



**PCT** WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro  
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<b>(51) Internationale Patentklassifikation <sup>6</sup>:</b> <b>C12N 15/31, C07K 14/37, C12N 1/21, A61K 39/35, G01N 33/569, C12N 15/62</b>		<b>A2</b>	<b>(11) Internationale Veröffentlichungsnummer: WO 95/06121</b> <b>(43) Internationales Veröffentlichungsdatum: 2. März 1995 (02.03.95)</b>
<b>(21) Internationales Aktenzeichen: PCT/AT94/00120</b> <b>(22) Internationales Anmeldedatum: 24. August 1994 (24.08.94)</b> <b>(30) Prioritätsdaten: A 1725/93 27. August 1993 (27.08.93) AT</b> <b>(71) Anmelder (für alle Bestimmungsstaaten ausser US): BIOMAY PRODUKTIONS- UND HANDELSGESELLSCHAFT M.B.H. [AT/AT]; Herrenstrasse 2, A-4020 Linz (AT).</b> <b>(72) Erfinder; und</b> <b>(75) Erfinder/Anmelder (nur für US): ACHATZ, Gernot [AT/AT]; Schießstattstraße 7/III/7, A-5020 Salzburg (AT). OBERKOFER, Hannes [AT/AT]; A-5732 Mühlbach 98 (AT). SIMON, Birgit [AT/AT]; Dirnböckweg 17, A-8700 Leoben (AT). UNGER, Andrea [AT/AT]; Zaisberg 14, A-5201 Seekirchen (AT). LECHENAUER, Erich [AT/AT]; Döttlstraße 16, A-5400 Hallein (AT). HIRSCHWEHR, Reinhold [AT/AT]; Nauseagasse 18/10, A-1160 Wien (AT). EBNER, Christoph [AT/AT]; St. Elisabethplatz 4/13, A-1040 Wien (AT). KRAFT, Dietric [AT/AT]; Montigasse 1, A-1170 Wien (AT). PRILLINGER, Hans-Jörg [AT/AT]; Ebersbrunn 70, A-3711 Ebersbrunn (AT). BREITENBACH, Michael [AT/AT]; Alfred Kubinstrasse 11/11, A-5020 Salzburg (AT).</b>		<b>(74) Anwälte: ITZE, Peter usw.; Amerlingstrasse 8, A-1061 Wien (AT).</b> <b>(81) Bestimmungsstaaten: AU, CA, FI, JP, NO, US, europäisches Patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</b> <b>Veröffentlicht</b> <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>	
<b>(54) Title: RECOMBINANT CLADOSPORIUM HERBARUM ALLERGENS</b> <b>(54) Bezeichnung: REKOMBINANTE CHLADOSPORIUM HERBARUM ALLERGENE</b> <b>(57) Abstract</b> <p>The invention relates to recombinant DNA molecules that code for polypeptides that have the antigenicity of the allergens Clah53, Clah47, Clah22 and Clah11 or for peptides that have at least one epitope of these allergens. These molecules are characterized in that they have nucleic acid sequences that homologously match the sequences, 1, 3-5, 7-9, 12-14 as well as 16 and 17, or with segments of these sequences, or have nucleic acid sequences that hybridize with the indicated nucleic acid sequences under stringent conditions.</p> <b>(57) Zusammenfassung</b> <p>Die Erfindung bezieht sich auf rekombinante DNA Moleküle, die für Polypeptide kodieren, die die Antigenität der Allergene Clah53, Clah47, Clah22 und Clah11 besitzen oder für Peptide, die mindestens ein Epitop dieser Allergene aufweisen. Diese Moleküle sind dadurch gekennzeichnet, daß sie Nukleinsäuresequenzen aufweisen, die mit den Sequenzen 1, 3-5, 7-9, 12-14 sowie 16 und 17, oder mit Teilbereichen dieser Sequenzen in homologer Weise übereinstimmen, bzw. Nucleinsäuresequenzen, die mit den genannten Nucleinsäuresequenzen unter stringenten Bedingungen hybridisieren.</p>			

RECEIVED

TECH CENTER 1600 2000

# **LEDIGLICH ZUR INFORMATION**

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AT	Österreich	GA	Gabon	MR	Mauretanien
AU	Australien	GB	Vereinigtes Königreich	MW	Malawi
BB	Barbados	GE	Georgien	NE	Niger
BE	Belgien	GN	Guinea	NL	Niederlande
BF	Burkina Faso	GR	Griechenland	NO	Norwegen
BG	Bulgarien	HU	Ungarn	NZ	Neuseeland
BJ	Benin	IE	Irland	PL	Polen
BR	Brasilien	IT	Italien	PT	Portugal
BY	Belarus	JP	Japan	RO	Rumänien
CA	Kanada	KE	Kenya	RU	Russische Föderation
CF	Zentrale Afrikanische Republik	KG	Kirgisistan	SD	Sudan
CG	Kongo	KP	Demokratische Volksrepublik Korea	SE	Schweden
CH	Schweiz	KR	Republik Korea	SI	Slowenien
CI	Côte d'Ivoire	KZ	Kasachstan	SK	Slowakei
CM	Kamerun	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Tschad
CS	Tschechoslowakei	LU	Luxemburg	TG	Togo
CZ	Tschechische Republik	LV	Lettland	TJ	Tadschikistan
DE	Deutschland	MC	Monaco	TT	Trinidad und Tobago
DK	Dänemark	MD	Republik Moldau	UA	Ukraine
ES	Spanien	MG	Madagaskar	US	Vereinigte Staaten von Amerika
FI	Finnland	ML	Mali	UZ	Usbekistan
FR	Frankreich	MN	Mongolei	VN	Vietnam

Rekombinante Cladosporium herbarum Allergene

Die Erfindung bezieht sich auf rekombinante DNA-Moleküle, die für Polypeptide kodieren, die die Antigenität der Allergene Clah 53, Clah47, Clah22 und Clah11 besitzen oder für Peptide, die mindestens ein Epitop dieser Allergene aufweisen.

Die genannten Allergene des extramuralen Schimmelpilzes Cladosporium herbarum, sowie auf die Molekülfragmente ( B- und T-Zell stimulierende Peptide) führen im Rahmen einer Immunantwort zu einem Überschießen der IgE Antikörperproduktion bei Pilzallergikern. Rekombinante Allergene bzw. immunogen wirkebde Teilpeptide können sowohl in vitro als auch in vivo zu einer verbesserten Diagnostik von Schimmelpilzallergien aber auch zur Induktion einer Immuntoleranz bzw. Anergie von allergenspezifischen T-Zellen herangezogen werden.

Das Immunsystem der Vertebraten entwickelte sich in der Evolution als wirksame Waffe gegen Angriffe auf das Individuum von außen, aber auch von innen. Unter "normalen" Bedingungen kann das Immunsystem zwischen "Selbst" und "Nicht-Selbst" unterscheiden. Wie man aber heute von vielen Regulationskaskaden weiß, laufen solche Regulationsmechanismen nicht immer fehlerlos ab, was im Falle der Immunologie einem Angriff auf eigenes Gewebe gleichkommt. Man kennt heute mehrere prinzipielle Situationen, in denen der Körper Opfer seines eigenen Immunsystems wird. Eine dieser unerwünschten immunologischen Antworten kann durch Umweltantigene ausgelöst werden. Die Reaktion, die durch diese Fehlsteuerung verursacht wird, nennt man Allergie oder Hypersensibilitäts-Reaktion. Gell und Coombs (1975) definierten 4 Hypersensibilitätstypen (Typ I,II,III und IV). Allergene wie Pilzsporen gehören zu den "Typ I" oder "Anaphylaxie-Hypersensibilitäts" auslösenden Antigenen.

Das typische Bild einer Typ I Hypersensibilität besteht darin, daß ein einmaliger Kontakt mit einem bestimmten Antigen (z.B. mit Pilzsporen) keinen nennenswerten Effekt auf das Individuum zeigt. Kommt es jedoch nach einigen Wochen ein zweites Mal zu einem Allergenkontakt, dann reagiert der nun sensibilisierte Organismus mit Symptomen einer allgemeinen Anaphylaxie. Eine Kontraktion der glatten Muskulatur sowie eine Dilatation der Kapillaren ist die Folge. Dies beruht darauf, daß der primäre Kontakt mit dem allergenen Proteine in einer ungewöhnlichen humoralen Immunantwort resultiert. Neben einer in diesem Fall erwünschten IgG-Antwort, antwortet der Allergiepatient mit einer massiven IgE-Produktion. Diese primär gebildeten IgEs, binden mit ihren Fc-Teilen an

spezifische hochaffine Fc-Epsilon-Rezeptoren, deren Hauptlokalisation an der Außenseite von Mastzellen und Basophilen zu finden ist. Bei einem Zweitkontakt mit dem Allergen kommt es zwischen den gebundenen IgE-Molekülen durch das Allergen zu einer Quervernetzung, was schließlich eine Degranulation der Mastzellen und Basophilen und somit die Freisetzung von Mediatorstoffen wie Histamin und 5 Arachidonsäuremetaboliten etc. zur Folge hat.

Die wichtigsten Umweltallergene sind Proteine mit einem Molekulargewichte zwischen 10 und 50kD. Die wichtigste Quelle dieser allergenen Proteine, die eine TypI-Allergie hervorrufen, stellen Inhalationsallergene wie Pilzsporen, Pollen, Kot von Hausstaubmilben etc. dar (Review Bold et al. 1973). Die für eine Pilzallergie relevanten Pilze gehören einer Gruppe von eukaryontischen, filamentös wachsenden, 10 Sporen bildenden Pilzen an. Da die Sporen die Verbreitungsform des Pilzes darstellen (die Sporen können durch den Wind leicht vertragen werden), ist anzunehmen, daß ihnen eine entscheidende Rolle bei der Allergieauslösung zukommt.

Man weiß heute, daß 20% der Atopiker durch Pilzsporen sensibilisiert wurden (Lazey 1981). Solche Patienten zeigen, wenn der obere Atmungstrakt mit 15 Pilzsporen in Kontakt gekommen ist, eine typische TypI-Allergie mit Symptomen wie Heuschnupfen und Asthma. Wenn Pilzsporen in einer solchen TypI-Antwort involviert sind, beträgt die Größe der Sporen mehr als 5µm. Die Größenselektion begünstigt Pilze wie *Cladosporium herbarum* und *Alternaria alternata* als Allergieauslöser.

*Cladosporium herbarum* (bzw. die Sporen von *Cladosporium herbarum*) ist 20 der am häufigsten vorkommende Pilz in der Luft (Gravesen 1979). Die sehr trockenen Sporen von *Cladosporium herbarum* können durch den Wind relativ leicht vertragen werden. In belasteten Zeiten ist es nicht selten, daß man an die 35000 Konidien pro m<sup>3</sup> Luft findet. Durch die leichte Verschleppung der Sporen ist an solchen Spitzentagen auch in geschlossenen Räumen eine erhöhte Sporenbelastung meßbar. Die Hauptbelastungszeit liegt zwischen Frühling und Frühherbst. Diese 25 hohe Konidienzahlen lassen sich damit erklären, daß *Cladosporium herbarum* wegen seiner "genügsamen" Lebensweise nahezu überall zu finden ist. Bevorzugte Lebensräume sind jedoch absterbende Pflanzen, verschiedene Bodentypen, aber auch diverseste Nahrungsmitteln. Nicht gereinigte Kühlschränke, Fensterrahmen, Strohdächer und verschiedene Textilien gehören zu den weiteren Stand- bzw. Lebensorten dieses Pilzes.

30 Aus diesen Gründen (ein Kontakt mit *Cladosporium herbarum* Sporen kann praktisch niemals gänzlich ausgeschlossen werden) ist es nicht verwunderlich, daß

Cladosporium herbarum zum Gegenstand intensiver allergologischer Forschung geworden ist. So reagieren zB. in Finnland 8% der asthmatischen Kinder positiv auf Cladosporium (Foucard et al. 1984).

Die Beschreibung der allergenen Proteine von Cladosporium herbarum erfolgt mittels zum Teil aufwendiger molekularbiologischen Techniken. Die vermutete Zahl  
5 von Cladosporium herbarum Allergenen liegt bei ca. 60 (Aukrust 1979, 1980). Das in der Literatur beschriebene Hauptallergen ClahI 1 wurde aus Rohextrakten aufgereinigt. Das Molekulargewicht liegt bei etwa 13kD. Klonierungen von diversen Cladosporium herbarum Allergenen sind noch nicht vorgenommen worden. Der Vorteil von gentechnologisch hergestellten allergenen Proteinen bzw. deren Teilpeptiden (Voraussetzung dafür ist jedoch eine immunologisch vergleichbare  
10 Reaktivität - konnte bei Betula verucosa (Ferreira et al. 1993) und anderen Allegenen schon gezeigt werden) liegt:

a) In der Verbesserung der Testsysteme wie RIA (Radioimmunassay), IRMA (Immunradiometrische Assay), ELISA (Enzyme-linked immunosorbent Assay), LIA (luminescence immunoassay), Immunblots, Histamine-release-assay, T-Zell Proliferationsassay und viele mehr.

15 b) In der Verbesserung der Hyposensibilisierungstherapie: Diese Therapie besteht in der Zufuhr von Allergenextrakten in Form von Injektionen oder peroraler Applikation in wässriger Form als Tropfen in steigender Dosierung, bis eine Erhaltungsdosis über mehrere Jahre erreicht ist. Resultat dieser Therapie ist das Erreichen einer Toleranz gegenüber den eingesetzten Allergenen, was sich in einer Abnahme der Krankheitssymptome äußert (Birkner et al. 1990). Das Problem bei  
20 dieser Art der Behandlung liegt in der Vielzahl der dadurch auftretenden Nebenwirkungen. Bei der Hyposensibilisierungstherapie sind Fälle von anaphylaktischem Schock während der Behandlung aufgetreten. Das Problem hierbei liegt in der schweren Standardisierbarkeit der Pilzprotein-Isolate. Bei einem Einsatz von von Allergenen abgeleiteten aber nicht anaphylaktisch wirkenden Peptiden könnten risikolos höhere Dosen verabreicht werden, wodurch eine wesentliche  
25 Verbesserung der Hyposensibilisierung erreicht werden kann.

c) Mit diesen Untersuchungen können aber auch spezifische T- und B-Zell-Epitope definiert werden. Solche Peptide besitzen die Fähigkeit zB. T-Lymphozyten zu stimulieren und zur Proliferation anzuregen, aber die Zellen (bei genau definierter Dosis) auch in einen Zustand der Toleranz bzw. Nicht-Reaktivität (Anergie) zu versetzen (Rothbard et al. 1991).

30 Erfindungsgemäß werden rekombinante DNA Moleküle der eingangs genannten Art geschaffen, die Nucleinsäuresequenzen aufweisen, die mit den

Sequenzen 1, 3-5, 7-9, 12-14 sowie 16 und 17, oder mit Teilbereichen dieser Sequenzen in homologer Weise übereinstimmen, bzw. Nucleinsäuresequenzen, die mit den genannten Sequenzen unter stringenten Bedingungen hybridisieren. Die DNA Moleküle können auch Nucleinsäuresequenzen aufweisen, die durch Degeneration von den vorgenannten Sequenzen ableitbar sind.

5 Weitere Merkmale gehen aus den nachstehenden Darlegungen hervor.

Beispiele:

a) Beschreibung der allergenen Proteine von *Cladosporium herbarum* mittels Western-Blotting

Für die Klonierung der vorliegenden Allergene von *Cladosporium herbarum* standen Sera von 142 Atopikern zur Verfügung. Um die Reaktivität der Patienten mit Pilzproteinextrakt zu testen, wurde *Cladosporium herbarum* (Sammlung Prof. 10 Windisch [Berlin] Nummer: 28-0202) in Flüssigmedium (2% Glukose, 2% Pepton, 1% Hefeextrakt) gezüchtet und anschließend lyophilisiert. Aus diesem Material wurden sodann die allergenen Proteine ausgewaschen und mittels Lyophilisator aufkonzentriert. Die Auftrennung erfolgte auf einem denaturierenden Polyacrylamidgel, das anschließend geblottet, mit Patientenserum inkubiert und mit 15 <sup>125</sup>I-markiertem anti human IgE detektiert wurde. In Prozentzahlen ausgedrückt reagierten die Patienten auf die allergenen Proteine wie folgt:

Clah53	53%
Clah47	53%
Clah22	8.2%
Clah11	4%

20 Wurde Protein aus gekauftem Pilzmaterial der Firma Allergon (Schweden) isoliert und für den Immunblot verwendet, konnte nahezu dasselbe Bandenmuster detektiert werden. Somit können Clah53 und Clah47 in Bezug auf das uns zur Verfügung stehende Patientenspektrum als Hauptallergene, Clah22 und Clah11 als Nebenallergene eingestuft werden.

Die angeschlossenen zwei Figuren zeigen einen Überblick über das zur 25 Klonierung der beschriebenen Allergene zur Verfügung gestandene Patientenspektrum. Das erste der beiden Bilder zeigt ein 13,5 %iges Acrylamidgel. Die Patienten mit der Nummer 19 und 35 (es handelt sich hier auch um die Patienten, die für das spätere Screenen verwendet wurden) zeigen Banden in der Größenordnung 53kD, 46kD und 22kD. Im zweiten Bild, es handelt sich hier um ein 17,5% Polyacrylamidgel, wird auch die kleine Molekulargewichtsbande (11kD) bei 30 Patient 35 sichtbar.

Fig.1 zeigt ein Westernblotting eines 13,5%iges Polyacrylamidgels nach Auftrennung von *Cladosporium herbarum* Proteinextrakt und Inkubation mit Sera verschiedener Patienten.

Fig.2 zeigt eine Auftrennung von *Cladosporium herbarum* Proteinextrakt auf einem 17,5%igen Polyacrylamidgel; Inkubation mit Patientensera; Detektion mit 5 Jod-markiertem anti-human IgE.

b) Konstruktion der cDNA Expressionsbank

Gesamt RNA wurde nach der sauren Guanidium-Phenol-Extraktionsmethode aus selbst gezüchtetem Pilzmaterial gewonnen. Poly(A)plus mRNA-Anreicherung  
10 erfolgte mit Oligo(dT) Cellulose der Firma Böhlinger. Die cDNA Synthese (1. und 2. Strang) wurde wie im Manual des Lambda ZAP-Systems der Firma Stratagene beschrieben durchgeführt. Die cDNA wurde anschließend (3'-seitig) mit *EcoRI* und (5'-seitig) mit *XbaI* Linkern versehen, in vorverdaute Lambda-ZAP-Arme ligiert und verpackt. Der Titer der Primärbank betrug 1000000 Klone.

15 c) Screening der cDNA Genbank mit Patientensera, in vivo Excision, Sequenzierung

Das Screenen der Expressionsbank erfolgte mittels Inkubation der "gelifteten" Phagenplaques mit einem Seragemisch aus 2 Patienten, von denen man durch das Westernblotting wußte, daß sie das Spektrum der detektierten Antigene abdecken. Die Detektion erfolgte wieder mit anti human IgE RAST Antikörper der Firma Pharmacia. Von den nach Sekundär- und Tertiärscreening übriggebliebenen 200  
20 positiven Klonen wurden 30 mit Hilfe eines Helferphagen in vivo exzidiert und zu einem bereits fertig sequenzierbaren Bluescriptvektor religiert (Durchführung wie im Manual des Lambda ZAP-Kits). Restriktionsverdaue der exzidierten Plasmide zeigten (*EcoRI*-*XbaI* Doppelverdaue) 4 verschiedene Inserttypen. Diese 4 Klone wurden nach der Sangermethode (Sanger 1977) sequenziert.

25 d) Expression der Clah53, Clah47, Clah22 und Clah11 cDNA's als

$\beta$ -Galaktosidasefusionsprotein

Mit Hilfe des vorher beschriebenen IgE-Screenings konnten vier vollständige cDNA-Klone erhalten werden. Die jeweiligen rekombinanten Plasmide wurden in den *E.coli* Stamm XL1-Blue transformiert und mit IPTG (Isopropyl- $\beta$ -D-thiogalactopyranosid) induziert. Der *E.coli* Gesamtproteinextrakt  
30 wurde anschließend elektrophoretisch aufgetrennt und auf Nitrozellulose geblottet. Das Fusionsprotein wurde mittels Serum IgE von Pilzallergikern und

einem mit einem  $^{125}\text{I}$ -markiertem Kaninchen-anti human IgE Antikörper (Pharmacia, Uppsala Schweden) detektiert.

Fig. 3 zeigt das rekombinante  $\beta$ -Galaktosidasefusionsprotein nach Inkubation mit Patientenserum und Detektion mit jodmarkiertem anti-human IgE. Der  $\beta$ -Galaktosidaseanteil des Fusionsproteins beträgt 36 Aminosäuren, was einem Molekulargewicht von 3800 Dalton gleichkommt. Unter Berücksichtigung dieser "Vergrößerung" des allergenen Proteins ist auch die Fig. 3 zu sehen. Spur 1 (Klon 1-1) und 4 (Klon 6-1) zeigen das rekombinante Fusionsprotein Clah47, jetzt um den Fusionsanteil größer. Spur 2 (Klon 3-2) zeigt das rekombinante Clah53 Allergen. Die rekombinanten Proteine von Clah22 und Clah11 sind auf dieser Figur nicht zu sehen.

10 Fig.3 zeigt somit eine Expression der rekombinanten Proteine Clah47 und Clah53 im Vektor BS-SK<sup>+</sup> nach IPTG Induktion.

#### e) Bestimmung von B- und T-Zell Epitopen bei den rekombinanten Allergenen

Die rekombinante Primärsequenz der Allergene bietet die Voraussetzung für die Vorhersage von B- und T-Zellepitopen mittels geeigneter Computerprogramme. Die bestimmten Epitope werden jeweils bei der Beschreibung des rekombinanten Proteins in eigenen Figuren angeführt. Mit diesen Untersuchungen können spezifische T- und B-Zell-Epitope definiert werden die die Fähigkeit besitzen zB. T-Lymphozyten zu stimulieren und zur Proliferation anzuregen, aber die Zellen (bei genau definierter Dosis) auch in einen Zustand der Toleranz bzw. Nicht-Reaktivität (Anergie) zu versetzen (Rothbard et al. 1991).

20 Die Suche nach B-Zellepitopen wurde mit Hilfe des GCG-Programmes (Genetics-Computer-Group) "PROTCALC", das jedoch von der Arbeitsgruppe um Prof. Modrow mit wesentlichen Parametern erweitert wurde, durchgeführt. Die Bestimmung beruht auf einer Abwägung der Parameter Hydrophilität (Kyte-Doolittle), Sekundärstruktur (Chou-Fasman), Oberflächenlokalisation (Robson-Garnier) und Flexibilität, wodurch die Antigenität von Teilpeptiden 25 errechnet wird.

Das Prinzip der T-Zellepitop-Voraussage erfolgte im Prinzip nach dem Algorithmus von Margalit et al. (1987). Das Prinzip besteht in der Suche nach amphipathischen Helices laut Primärsequenz des zu bestimmenden Proteins, flankiert von hydrophilen Bereichen. Der berechnete Score muß für relevante T-Zellepitope größer als 10 sein. Bei MHC II (major histo compatibility locus) assoziierten 30 Peptiden kann kein Konsensus, weder der Sequenz noch der Länge des Peptids nach, wie bei HLA-A2 (human leucocyte antigen) assoziierten definiert werden. Bei



HLA-A2 assoziierten Peptiden beträgt die Länge des Peptids 10 Aminosäuren, wobei die 2. Aminosäure ein Tyrosin und die letzte Aminosäure ein Leucin darstellt (Rammensee et al. 1993). Die berechneten Epitope werden bei der Beschreibung der einzelnen allergenen Sequenzen getrennt angeführt.

## 5 Molekulare Charakterisierung der klonierten Pilzallergene (Sequenzprotokolle)

Im folgenden Kapitel werden nun die cDNA Sequenzen und die mit ihnen durchgeführten Analysen der Reihe nach angeführt. Die Computerauswertung der nachfolgenden Sequenzen wurden auf einer Ultrix-DEC 5000 Workstation unter Zuhilfenahme des GCG-Softwarepaketes (=Wisconsin Paket: die Algorithmen dieses 10 Paketes wurden von der Universität Wisconsin entwickelt) durchgeführt.

### A. Clah53

Die nachfolgende Sequenz 1 zeigt die vollständige cDNA Sequenz und die daraus abgeleitete Aminosäuresequenz, beginnend mit dem Start-Methionin. Das berechnete Molekulargewicht beträgt 53364 Dalton und entspricht somit der Größe 15 nach dem im Westernblot detektierten allergenen Protein in der Größe von 53kD. Dem reifen Protein dürfte nach bisheriger Analyse kein Signalpeptid voranstehen. Der offene Leserahmen von Clah53 beträgt 1491 Basenpaare bzw. 497 Aminosäuren.

Sequenz 1: Clah53=ALDH\_clado -> 1-phase Translation 53364 Dalton

## 20 (1) ANGABEN ZU SEQ ID NO:1

### (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 1491 Basenpaare / 497 Aminosäurereste

(B) ART: Nukleinsäure / Protein

(C) STRANGFORM: ds

## 25 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA zu mRNA / Protein

(iii) HYPOTHETISCH: nein

(iv) ANTISENSE: nein

(v) ART DES FRAGMENTS: Gesamtsequenz

### (vi) URSPRÜNGLICHE HERKUNFT:

## 30 (A) ORGANISMUS: Cladosporium herbarum

## (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

DNA Sequenz 1491 b.p. ATGACTTCCGTC ... CTCTTCGGTTAG linear

```

1   /   1
ATG ACT TCC GTC CAG CTT GAG ACG CCG CAC TCC GGC AAG TAC GAG CAG CCG ACC GGC CTC
met thr ser val gln leu glu thr pro his ser gly lys tyr glu gln pro thr gly leu
61   /   21
5 TTC ATC AAC AAC GAG TTC GTC AAG GGC CAA GAA GGC AAG ACC TTC GAT GTC ATC AAC CCC
phe ile asn asn glu phe val lys gly gln glu gly lys thr phe asp val ile asn pro
121   /   41
TCC GAC GAG AGC GTG ATC ACC CAG GTC CAC GAG GCC ACC GAG AAG GAT GTC GAC ATC GCC
ser asp glu ser val ile thr gln val his glu ala thr glu lys asp val asp ile ala
181   /   61
Gtc GCC GCC GCC CGc CAA GCC TTC GAG GGC TCA tgg aga ctg gag aca CCC GAG AAC CGT
val ala ala ala arg gln ala phe glu gly ser trp arg leu glu thr pro glu asn arg
241   /   81
GGC AAG CTG CTC AAC AAC CTC GCC AAC CTG TTC GAG AAG AAC ACT GAC CTC CTT GCT GCC
gly lys leu leu asn asn leu ala asn leu phe glu lys asn thr asp leu leu ala ala
301   /   101
10 GTT GAG TCG CTC GAC AAC GGC AAG GCC ACT TCC ATG GCA AGG GtG ACA tca GCA TGT GCG
val glu ser leu asp asn gly lys ala thr ser met ala arg val thr ser ala cys ala
361   /   121
TCC GGC TGC CTC AGA TAC TAC GGT GGT TGG GCG GAC AAG ATC ACC GGC AAG GTC ATC GAC
ser gly cys leu arg tyr tyr gly gly trp ala asp lys ile thr gly lys val ile asp
421   /   141
ACT ACG CCC GAC ACT TTC AAC TAC gTC AAG AAG GAG CCC ATT GGT GTT TGC cgg TCA GAT
thr thr pro asp thr phe asn tyr val lys lys glu pro ile gly val cys arg ser asp
481   /   161
CAT tCC CTG GAA CTT CCC CTT CTC ATG TGG GCA TGG AAG ATC GGC CCG GCC ATT GCT TGC
his ser leu glu leu pro leu leu met trp ala trp lys ile gly pro ala ile ala cys
541   /   181
15 GGT AAC ACT GTC Gtc CTG AAG ACT GCT GAG CAG ACC CCT CTT GGT GGT CTC GTC GCT GCC
gly asn thr val val leu lys thr ala glu gln thr pro leu gly gly leu val ala ala
601   /   201
AGC CTC GTC AAG GAG GCC GGT TTC CCT CCT GGTGGTC ATC AAC GTC ATT TCC GGT TTC GGC
ser leu val lys glu ala gly phe pro pro gly val ile asn val ile ser gly phe gly
661   /   221
AAG GTC GCT GGT GCC GCT CTC TCT TCT CAC ATG GAC GTC GAC AAG GTg GCC TTC ACC GGT
lys val ala gly ala ala leu ser ser his met asp val asp lys val ala phe thr gly
721   /   241
TCC ACC GTt GTC GGC cGC ACA ATC CTC AAG GCT GCT GcC tct TCC AAC TTG AAG AAG GTC
ser thr val val gly arg thr ile leu lys ala ala ala ser ser asn leu lys lys val
781   /   261
20 ACC cTC GAG CTC GGT GGC AAG TCA CCA AAC ATT GTC TTC GAG GAC GCC GAT ATT GAC AAC
thr leu glu leu gly gly lys ser pro asn ile val phe glu asp ala asp ile asp asn
841   /   281
GCC ATC TCA TGG GTC AAC TTC GGT ATC TTC TTC AAC CAC GGC CAG TGC TGC TGT GCT GGT
ala ile ser trp val asn phe gly ile phe phe asn his gly gln cys cys cys ala gly
901   /   301
TCG CGT GTG TAC GTT CAG GAG AGC ATC TAC GAC AAG TTC GTC CAG AAG TTC AAG GAG CGC
ser arg val tyr val gln glu ser ile tyr asp lys phe val gln lys phe lys glu arg
961   /   321
GCA CAG AAG AAC GTT GTT GGC GAC CCC TTC gcC GCC GAC ACA TTC CAG GGT CCT CAG GTT
ala gln lys asn val val gly asp pro phe ala ala asp thr phe gln gly pro gln val
1021   /   341
25 TCC AAG GTT CAG TTC GAC CGC ATT ATG GAG TAC ATC CAG GCC GGC AAG GAC gcc gGT GCC
ser lys val gln phe asp arg ile met glu tyr ile gln ala gly lys asp ala gly ala
1081   /   361
ACC GTC GAG ACC GGT GGA AGC cGT AAG GGT GAC AAG GGC TAC TTC ATT GAG CCC ACC ATC
thr val glu thr gly gly ser arg lys gly asp lys gly tyr phe ile glu pro thr ile
1141   /   381
TTC TCC AAC GTC ACA GAG GAC ATG AAG ATC GTG AAA GAG GAG ATC TTC GGC CCC GTC Tgc
phe ser asn val thr glu asp met lys ile val lys glu glu ile phe gly pro val cys
1201   /   401
tCG ATC GCC AAG TTC AAG ACC AAG GAG GAC GCC ATC AAG CTC ggc AAC GCC AGC ACA TAC
ser ile ala lys phe lys thr lys glu asp ala ile lys leu gly asn ala ser thr tyr
1261   /   421
30 GGT CTC GCC GCC GCC GTC CAC ACC AAG AAC CTC AAC ACC GCC ATC GAG GTC TCC AAC GCT
gly leu ala ala ala val his thr lys asn leu asn thr ala ile glu val ser asn ala
1321   /   441
CTC AAG GCG GGC ACC GTC TGG GTC AAC ACT TAC AAC ACC CTC CAC CAG ATG CCg TTC

```

```

leu lys ala gly thr val trp val asn thr tyr asn thr leu his his gln met pro phe
1381 / 461                               1411 / 471
GGT GGC TAC AAG GAG TCC GGT ATT GGC CGT GAG CTC GGC GAG GAT GCG TTG GCC AAC TAC
gly gly tyr lys glu ser gly ile gly arg glu leu gly glu asp ala leu ala asn tyr
1441 / 481                               1471 / 491
ACT CAG ACC AAG ACT GTT AGC ATT CGT CTC GGT GAC GCC CTC TTC GGT TAG
thr gln thr lys thr val ser ile arg leu gly asp ala leu phe gly amb

```

- 5 Eine Homologiesuche der gezeigten Proteinsequenz in der SWISSPROT-Proteindatenbank zeigte signifikante Homologie des Proteins zu verschiedenen Aldehyddehydrogenasen. Das nachfolgende "multiple Alignment" mit ALDH-Sequenzen von Aspergillus, Rind, Pferd, Maus, Ratte, Mensch, E.coli. Pseudomonas sowie den Pilzen Cladosporium herbarum und Alternaria alternata spiegelt die hohe Homologie von Clah53 zu Aldehyddehydrogenasen wieder (Seq.2).
- 10 Der ermittelte Konsensus zeigt Identitäten von Aminosäuren quer durch alle Organismen.

## Sequenz 2:

```

1                               50
pileup.msf(ALDH_alt)          .....M tsvklstpq t g..EFE.Qpt glFINNEFvk
15 pileup.msf(ALDH_clado)      .....M tsvqlstphs g..kYE.Qpt glFINNEFvk
pileup.msf(ALDH_asp)          .....msdlfttiet PvikYE.Qpl glFINNEFvk
pileup.msf(dham_bovin)        mlravalaaa xlgprqgrxL lSaAtqaVpt PnqqpEVlyn qIFINNEWHd
pileup.msf(dham_horse)        .....aaaAtqaVpa PnqqpEVfyn qIFINNEWHd
pileup.msf(dhac_mouse)        .....SSpaqPrVpa PLAdLkIQht KIFINNEWHn
pileup.msf(dhac_rat)          .....SSpaqPaVpa PLanLkIQht KIFINNEWHn
pileup.msf(dhac_human)        .....SSagtPdIPv lLtDLkIQyT KIFINNEWHd
pileup.msf(dhab_ecoli)        .....srmacq qlYIhggyts
pileup.msf(dmpe_psepu)        .....mkel KhFINGaFvg
Consensus                    -----I-----

51                               100
20 pileup.msf(ALDH_alt)        aVdGKTFdVI NPstEEVics VqEatekDVD lAVaAARKAF n..gpWaket
pileup.msf(ALDH_clado)        gqeGKTFdVI NPsdEsVItq VhEatekDVD lAVaAARQAF e..gsWRlet
pileup.msf(ALDH_asp)          gVeGKTFqVI NPsnEkVits VhEatekDVD vAVaAARaAF e..gpWRqvt
pileup.msf(dham_bovin)        aVskKTFpTv NPstGdVICH VaEgDKADVD xAVKAARaAF QlGSPWRrMD
pileup.msf(dham_horse)        aVskKTFpTv NPstGdVICH VaagDKADVD xAVKAARaAF QlGSPWRrMD
pileup.msf(dhac_mouse)        sVSGKkFFvI NPAtEEVICH VeEgDKADVD xAVKAARQAF QIGSPWRrMD
pileup.msf(dhac_rat)          slngKkFFvI NPAtEEVICH VeEgDKADVD xAVKAARQAF QIGSPWRrMD
pileup.msf(dhac_human)        sVSGKkFFvI NPAtEEelCq VeEgDKADVD xAVKAARQAF QIGSPWRrMD
pileup.msf(dhab_ecoli)        atSGrTFeti NPangNVlat VqaagreDVD xAVKsAQOgq k...iWasnt
pileup.msf(dmpe_psepu)        saSGrTFedv NPangqVlar VhEagrAEVD aAVqAARaAL k..gpWgkMs
Consensus                    -----F--- NP----- V-----VD -AV--A---- -----W----

101                               150
pileup.msf(ALDH_alt)          penRGkLLNK LADLFeknaD LiAAvEaLDN GKAFsMAknV DvpaaagCLR
pileup.msf(ALDH_clado)        penRGkLLNn lAnLFekntD LLAAVESLDN GKATsMAR.V tsacaagCLR
pileup.msf(ALDH_asp)          pSERGiLLNK LADLMErdID tLAAIESLDN GKAFtMAK.V DLansIgCLR
pileup.msf(dham_bovin)        ASERGRLlNr LADLIERDRt yLAALEtLDN GKpYiisYLV DLdmvLKCLR
pileup.msf(dham_horse)        ASDRGRLlNr LADLIERDRt yLAALEtLDN GKpYvisYLV DLdmvLKCLR
pileup.msf(dhac_mouse)        ASERGcLLNK LADLMErdRI LLAtMEaIng GKvFanAYLs DLggcIKaIk
pileup.msf(dhac_rat)          ASERGcLLNK LADLMErdRV LLAtMESMna GKlFthAYLI DtevsIKaIk
pileup.msf(dhac_human)        ASERGRLlYK LADLIERDRl LLAtMESMng GKlYSnAYLn DLageIKtLR
pileup.msf(dhab_ecoli)        AmERsRilrr avDiLrernD eLakLEtLDt GKAYSetstV DivtgadvLe
30 pileup.msf(dmpe_psepu)        vSERaeiLhr vADgitarFD eFleaEcLDt GKpkSLashI DiprgaanFk
Consensus                    ---R----- -----E---- GK-----

```

```

151                                     200
pileup.msf(ALDH_alt) YYgGW...A DKIEGKVvdt apDsFnYiRK .slivfavrs smelPLlMWs
pileup.msf(ALDH_clado) YYgGW...A DKITGKvIdt tpdTfNvYkK EPiGVCrsdh slePLlMWa
pileup.msf(ALDH_asp) YYAGW...A DKIHGqTIdt npEtLTYYTRh EPVGVCgQII PWNFFlLMWs
pileup.msf(dham_bovin) YYAGW...A DKyHGKTIPi DGDYFsYTRh EPVGVCgQII PWNFFlLMqa
pileup.msf(dham_horse) YYAGW...A DKyHGKTIPi DGDFFsYTRh EPVGVCgQII PWNFFlLMqa
pileup.msf(dhac_mouse) YCAGW...A DKIHGqTIPs DGDIFTYYTRr EPiGVCgQII PWNFFlLMfI
pileup.msf(dhac_rat) YFAGW...A DKIHGqTIPs DGDvFTYYTRr EPiGVCgQII PWNgPLlLFi
pileup.msf(dhac_human) YCAGW...A DKiGqRTIPi DGNfFTYYTRh EPiGVCgQII PWNFFlVMLi
pileup.msf(dhab_ecoli) YYAGli..pA legsqiplre ts..FvYTRr EPiGVvagig aWNYPIqial
5 pileup.msf(dmpe_psepu) VFADllknvA teafematPd gsgainYavr rPkGVivGis PWNlPLlLmt
Consensus -----A -----Y-----V-----P-----

```

```

201                                     250
pileup.msf(ALDH_alt) WKIGPAIATG NTVVLKtAEQ TPLsAyiack LIqZAGFFPG VINVITGFGK
pileup.msf(ALDH_clado) WKIGPAIACG NTVVLKtAEQ TPLggLvAAS LVKEAGFFPG VINVISGFGK
pileup.msf(ALDH_asp) WKIGPAvAaG NTVVLKtAQq TPLsALYaaK LIKEapFPaG VINVISGFGGr
pileup.msf(dham_bovin) WKIGPALATG NvVVMKVAEQ TPLTALYvAn LIKEAGFFPG VVNVPFGFGP
pileup.msf(dham_horse) aKlGPALATG NvVVMKVAEQ TPLTALYvAn LtKEAGFFPG VVNVPFGFGP
pileup.msf(dhac_mouse) WKIGPALsCG NTVVVKPAEQ TPLTALhLAS LIKEAGFFPG VVNVPFGYGP
pileup.msf(dhac_rat) WKIGaALsCG NTVVVKPAEQ TPLTALYMAS LIKEAGFFPG VVNVPFGYGP
pileup.msf(dhac_human) WKIGPALsCG NTVVVKPAEQ TPLTALhVAS LIKEAGFFPG VVNVPFGYGP
10 pileup.msf(dhab_ecoli) WKsaPALAaG NamIFKPsEv TPLTALKLaE IySEAGLPdG VNVlPVGaG
pileup.msf(dmpe_psepu) WKVGPALACG NTVVVKPsEe TPLTtallGe vmqaAGvPaG VynVvhGFGP
Consensus -K---A---G N---K--- TPL-----A---P-G V-N---G-G-

```

```

251                                     300
pileup.msf(ALDH_alt) .iAGAAmSaH MDIDKIAFTG STvVGrqIMx sAagsNLKkv TLELGGKSPN
pileup.msf(ALDH_clado) .VAGAAISSH MDVDKVAFTG STvVGrtilK AAasNLKkv TLELGGKSPN
pileup.msf(ALDH_asp) .TAGAAISSH MDIDKVAFTG STlVoptILq AAaKSNLkv TLELGGKSPN
15 pileup.msf(dham_bovin) .TAGAAIASH eDVdKVAFTG STEVghLIqv AAgKSNLkv TLElGGKSPN
pileup.msf(dham_horse) .TAGAAIASH eDVdKVAFTG STEVghLIqv AAgKSNLkv TLELGGKSPN
pileup.msf(dhac_mouse) .TAGAAISSH MDVDKVAFTG STqVgKLIkE AAgKSNLkv TLELGGKSPc
pileup.msf(dhac_rat) .TAGAAISSH MDIDKVsFTG STEVgKLIkE AAgKSNLkv TLELGGKSPc
pileup.msf(dhac_human) .TAGAAISSH MDIDKVAFTG STEVgKLIkE AAgKSNLkv TLELGGKSPc
pileup.msf(dhab_ecoli) E.tGqylteH pgIakVsFTG gvasGKkvMa nsaaSsLkEv TMELGGKSPl
pileup.msf(dmpe_psepu) DsAGAfiteH pDVnaItFTG eTtGeaIM. rAaakgvrvP sFELGGKnaG
Consensus ---G-----H -----FTG ----G-----V ---E-GGK---

```

```

301                                     350
pileup.msf(ALDH_alt) IVFADADLDe AIhWvNFGIY FNhGQaCCAG SRIYVQeEiY DkFIQRfKER
pileup.msf(ALDH_clado) IVFeDADIDN AIhWvNFGIF FNhGQCCcAG SRVYVQESiY DkFVQkfKER
20 pileup.msf(ALDH_asp) IVFdDADIDN AISWanFGIF FNhGQCCcAG SRIYVQeGiY DkFVaRfKER
pileup.msf(dham_bovin) IImSDADMDw AVEqAHFaIf FNQGCCcAG SRIYVQeDiY aEFvRSVaR
pileup.msf(dham_horse) IIvsDADMDw AVEqAHFaIf FNQGCCcAG SRIYVQeDvY aEFvRSVaR
pileup.msf(dhac_mouse) IVFADADLDI AVEFAHhGVF YhQGQCCvAa SRIYVQeSiY DEFVrSVrER
pileup.msf(dhac_rat) IVFADADLDs AVEFAHhGVF FhQGQICvAa SRIYVQeSiY DEFVrSVrER
pileup.msf(dhac_human) IVLADADLDN AVEFAHhGVF YhQGQCCIAa SRIYVQeSiY DEFVrSVrER
pileup.msf(dhab_ecoli) IVFdDADLDI AadIAmmanF FssGQvCtnG tRVFvpakck aaFeQkilaR
pileup.msf(dmpe_psepu) IVFADcDLDk AIEgsmrsvF aNgGQvClgt eRIYVerpIF DEFVaRlkag
Consensus I---D-D-D- A-----GQ-C---R--V-----F-----

```

```

351                                     400
25 pileup.msf(ALDH_alt) AaqnaVGdPF aa.TlQGpQV sqLQFDriMg YIEgKksGA tiZtGGnRkG
pileup.msf(ALDH_clado) AqKnVVGdPF aadTfQGpQV sKVQFDriME YIqaGKdaGA tvZtGGsRkG
pileup.msf(ALDH_asp) AqKnkvGNPF EqdTfQGpQV sqLQFDriME YInhgKkaGA tvatGGdRhG
pileup.msf(dham_bovin) AKsrVVGnPF DsrTeQGpQV detQFkKVLg YIkSGKeEGl KllCGGgaaa
pileup.msf(dham_horse) AKsrVVGnPF DsqTeQGpQV detQFkKVLg YIkSGKeEGA KllCGGgaaa
pileup.msf(dhac_mouse) AKKyVlGNPL tPgInQGpQI dKEQhKILD lIESGKkEGA KlECGGGRWG
pileup.msf(dhac_rat) AKKyVlGNPL DsgIsQGpQI dKEQhKILD lIESGKkEGA KlECGGGRWG
pileup.msf(dhac_human) AKKyIlGNPL tPgvtQGpQI dKEQYDKILD lIESGKkEGA KlECGGgPWG
pileup.msf(dhab_ecoli) verIraGdvF DPqTnfGPlV sfphrDnVLr YIakGKeGA rvlCGGdvIk
pileup.msf(dmpe_psepu) AeslVIGtPd DPqanfGPII alqhrEKVLs YyqkavDEGA tvvtGGGvpe
Consensus -----G--- -----GP--- -----G-----GG-----

```

```

      pileup.msf(ALDH_alt)      .....DKGYF IePTIFSNTV EDMkIqqEEI FGPVctIsKF KtkADVikig
      pileup.msf(ALDH_clado)    .....DKGYF IePTIFSNTV EDMkIvKEEI FGPVcsIaKF KtkEDaIKlg
      pileup.msf(ALDH_asp)      .....neGYF IQPTVFtdVT sDMkIAqEEI FGPVvtIqKF KdVaEaIKig
      pileup.msf(dham_bovin)    .....DrGYF IQPTVFgdlq DgMtIAKEEI FGPVmqILKF KSmEEVVGRA
      pileup.msf(dham_horse)    .....DrGYF IQPTVFgdlq DgMtIAKEEI FGPVmqILKF KtIEEVVGRA
      pileup.msf(dhac_mouse)    .....nKGFF VQPTVFSNTV DEMrIAKEEI FGPVqQIMKF KSVDDVIKRA
      pileup.msf(dhac_rat)      .....nKGFF VQPTVFSNTV DEMrIAKEEI FGPVqQIMKF KSIDDEVIKRA
      pileup.msf(dhac_human)    .....nKGFF VQPTVFSNTV DEMrIAKEEI FGPVqQIMKF KSIDDEVIKRA
      pileup.msf(dhac_ecoli)    .gdgFDnGaw VaPTVFtdcs DDmtIvrEEI FGPVmsILTY eSeDEVIRRA
      pileup.msf(dmpe_psepu)    mpaeLagGaw VQPTIWTgla DgaaVvtEEI FGPcchIrfp dreEEaVela
5      Consensus      -----G-- --PT----- -----EEI FGP---I--- -----

```

```

      pileup.msf(ALDH_alt)      451
      pileup.msf(ALDH_clado)    NNTTYGLsAA VhtsnLttAI eVanALrAGT VWNsYntLh wQlPFGGYKe
      pileup.msf(ALDH_asp)      NstTYGLAAA VhtKnLntAI eVsnALrAGT VWNtYntLh hQmPFGGYKe
      pileup.msf(dham_bovin)    NnsKYGLAAA VFTKDLDKAn ylsqALQAGT VWNcydvFg AQsPFGGYKL
      pileup.msf(dham_horse)    NnsKYGLAAA VFTKDLDKAn ylsqALQAGT VWINcydvFg AQsPFGGYKL
      pileup.msf(dhac_mouse)    NNTTYGLAAG lFTKDLDKAI tVssALQAGT VWNcyimLs AQcPFGGFKM
      pileup.msf(dhac_rat)      NNTpYGLAAG VFTKDLDRAI tVssALQAGT VWNcyimLs AQcPFGGFKM
      pileup.msf(dhac_human)    NNTfYGLsAG VFTKDIDKAI tISSALQAGT VWNcygVvs AQcPFGGFKM
      pileup.msf(dhac_ecoli)    NdTDYGLAAG lVtADLnRAh rVihqLeAGi cWINTWgesp AempVGGYKh
10     pileup.msf(dmpe_psepu)    NslpYGLAAt lWTentsrAh rVagqLeAGi VWNsWflrd lrtafGGsKq
      Consensus      N---YGL-A- --T-----A- -----L-AG- -W-N----- -GG-K-

```

```

      pileup.msf(ALDH_alt)      501
      pileup.msf(ALDH_clado)    SGiGRELGEa aLdnYiqTKT VsIrlgdvlf GZ
      pileup.msf(ALDH_asp)      SGiGRELGEa aLanYTqTKT VsIrlgdalf GZ
      pileup.msf(dham_bovin)    SGiGRELGSY aLenYTqTKT Vhyrlgdalf a.
      pileup.msf(dham_horse)    SGsGRELGEY GLqayTEVKT VTVrvpQKNS ..
      pileup.msf(dhac_mouse)    SGnGRELGEY GLqayTEVKT VTIKvpQKNS ..
      pileup.msf(dhac_rat)      SGnGRELGEh GLyeYTELKT VamKisQKNS ..
      pileup.msf(dhac_human)    SGnGRELGEY GvyeYTELKT VamKisQKNS ..
15     pileup.msf(dhac_ecoli)    SGnGRELGEY GFheYTEVKT VTVKisQKNS ..
      pileup.msf(dmpe_psepu)    SGiGREngvm tLqsYTqVKs IqVemakfqs if
      Consensus      SGiGREgVh sLefYTELKn IcVKL..... ..

```

Die NAD-abhängige ALDH ist das Hauptenzym, das an der Oxidation von Azetaldehyd, ein Primärprodukt des Alkoholmetabolismus, im Menschen beteiligt ist. Isoenzyme sind hierbei oft zu finden (Harada et al. 1982). Beim Menschen z.B. findet man das Isoenzym ALDH I in Mitochondrien, ALDH II im Zytoplasma. Interessanterweise ist die Abwesenheit von ALDH I bei Asiaten keine Seltenheit (Harada et al. 1982). Die Defizienz von ALDH I resultiert in einem hohen Spiegel von Acetaldehyd, was sich als sogenanntes "flushing syndrome", sowie anderen vasomotorischen Symptomen nach Alkoholgenuß bemerkbar macht. Der Isoenzymverlust läßt sich auf eine Mutation zurückführen, die das native Protein in seiner Struktur verändert (Hsu et al. 1987). Der Zusammenhang zwischen ALDH und Allergieauslösung ist zum Zeitpunkt noch nicht bekannt.

Die nachfolgende Sequenz 3 zeigt die mit Computersuche identifizierten Bereiche mit hohem antigenem Index. Diese Bereiche stellen hochpotente B-Zellepitope dar.

Sequenz 3: Clah53=ALDH\_clado: B-Zellepitope

5 (1) ANGABEN ZU SEQ ID NO:3

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: einzeln angeführt

(B) ART: Protein

(ii) ART DES MOLEKÜLS: Peptide

10 (iii) HYPOTHETISCH: nein

(v) ART DES FRAGMENTS: N-Terminus bis C-Terminus

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Cladosporium herbarum

(C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

15

Ser Val Gln Leu Glu Thr Pro His Ser Gly Lys Tyr Glu Gln Pro Thr Gly (3-19)

Asn Asn Glu Phe Val Lys Gly Gln Glu Gly Lys Thr Phe (23-35)

Ile Asn Pro Ser Asp Glu Ser Val Ile (38-46)

Ala Ala Ala Arg Gln Ala Phe Glu Gly Ser Arg Lys Glu Thr Pro Glu Asn Arg Gly Lys Leu Leu  
Asn (62-84)

20 Cys Ala Ser Gly Cys Leu Arg Tyr Tyr Gly Gly Trp Ala Asp Lys Ile Thr Gly (118-135)

Lys Val Ile Asp Thr Thr Pro Asp Thr Phe Asn Tyr Ser Arg Arg Ser Pro Leu Val (136-154)

Leu Lys Thr Ala Glu Gln Thr Pro Leu GlyGGly (184-194)

Ala Ser Leu Val Lys Glu Ala Gly Phe Pro Pro Gly Val Ile Asn (198-212)

Ala Ala Ser Ser Asn Leu Lys Lys Val Thr Leu (250-260)

Glu Leu Gly Gly Lys Ser Pro Asn Ile Val Phe (261-271)

25 Tyr Asp Lys Phe Val Gln Lys Phe Lys Glu Arg Ala Gln Lys Asn Val Val Gly (308-325)

Tyr Ile Gln Ala Gly Lys Asp Ala Pro Ser Thr Val Glu Thr Gly Gly Ser Gly Lys Gly Asp Lys  
Gly Tyr Phe Ile (349-374)

Ile Ala Lys Phe Lys Thr Lys Glu Asp Ala Ile Lys Leu Asn Ala Ser Thr Tyr Gly Leu Ala  
(400-420)

Trp Val Asn Thr Tyr Asn Thr Leu His His Gln Met Pro Phe Gly Gly Tyr Lys Glu Ser Gly Ile

30 Gly Arg Glu Leu Gly Glu (444-471)

Asp Ala Leu Ala Asn Tyr Thr Gln Thr Lys Thr Val Ser Ile (472-485)

Die nachfolgende Sequenz 4 zeigt die mit Hilfe des Computerprogrammes bestimmten amphipathischen Helices, die von hydrophilen Bereichen flankiert werden. Solche Bereiche, mit einem Score höher als 10, stellen mögliche T-Zellepitope dar.

5

#### Sequenz 4: Vorausgesagte amphipathatische Segmente

##### T-Zellepitope

-----  
 PHSGKYE  
 KTFDVIN  
 10 KLLNNLANLFE  
 AAVESLDNGKATS  
 GCLRYYGWADKITGKVIDTTP  
 GVINVISGFGKVAGAAL  
 IYDKFVQKFKERAQKNV  
 QFDRIMEYIQA  
 APSTVETGGSG  
 15 FSNVTEEM  
 EVSNALK  
 NTYNTL  
 RELGEDALANYTQT

#### (1) ANGABEN ZU SEQ ID NO:4

20

##### (i) SEQUENZKENNZEICHEN:

(A) LÄNGE: einzeln angeführt

(B) ART: Protein

(ii) ART DES MOLEKÜLS: Peptide

(iii) HYPOTHETISCH: nein

25 (v) ART DES FRAGMENTS: N-Terminus bis C-Terminus

##### (vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Cladosporium herbarum

(C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

Pro His Ser Gly Leu Lys Tyr Glu (9-15)

Lys Thr Phe Asp Val Ile Asn (27-31)

30

Lys Leu Leu Asn Asn Leu Ala Asn Leu Phe Glu (81-91)

Ala Ala Val Glu Ser Leu Asp Asn Gly Lys Ala Thr Ser (98-110)  
 Gly Cys Leu Arg Tyr Tyr Gly Gly Trp Ala Asp Lys Ile Thr Gly Lys Val Ile Asp Thr Thr Pro  
 (121-142)  
 Gly Val Ile Asn Val Ile Ser Gly Phe Gly Lys Val Ala Gly Ala Ala Leu (209-225)  
 Ile Tyr Asp Lys Phe Val Gln Lys Phe Lys Glu Arg Ala Gln Lys Asn Val (307-323)  
 5 Gln Phe Asp Arg Ile Met Glu Tyr Ile Gln Ala (342-352)  
 Ala Pro Ser Thr Val Glu Thr Gly Gly Ser Gly (356-366)  
 Phe Ser Asn Val Thr Glu Glu Met (379-386)  
 Glu Val Ser Asn Ala Leu Lys (433-439)  
 Asn Thr Tyr Asn Thr Leu (446-451)  
 Arg Glu Leu Gly Glu Asp Ala Leu Ala Asn Tyr Thr Gln Thr (467-480)

10

Die T-Zellepitope errechnen sich aus den Aminosäurepositionen der Midpoints, die N-terminal von einem Lysin (K), C-terminal von einem Prolin (P) flankiert werden (=Flags). Es sind nur dann potentielle T-Zellepitope vorhanden, wenn der "Score-Index" größer als 10 ist.

#### 15 B. Clah47

Die nachfolgende Sequenz 5 zeigt die vollständige cDNA Sequenz des allergenen Proteins Clah47. Die Aminosäuresequenz wurde von der DNA Sequenz hergeleitet. Auch bei diesem Protein sind keine Anzeichen einer Signalsequenz vorhanden. Die Gesamt DNA Sequenz beträgt 1323 Basenpaare, was einer Proteinlänge von 441 Aminosäuren entspricht. Das berechnete Molekulargewicht des  
 20 rekombinanten Proteins beträgt 47617 Dalton und entspricht somit der detektierten Bande (47kD) im Westernblot. Wie zu Beginn erwähnt, wird das allergene Protein mit dem Molekulargewicht von 47kD von 53% der Patienten erkannt und stellt somit ein wichtiges Hauptallergen dar.

25 Sequenz 5: Clah47=Enolase\_clado -> 1-phasen Translation 47617 Dalton

(1) ANGABEN ZU SEQ ID NO:5

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 1323 Basenpaare / 441 Aminosäurereste

(B) ART: Nukleinsäure / Protein

30 (C) STRANGFORM: ds



- (D) TOPOLOGIE: linear  
(ii) ART DES MOLEKÜLS: cDNA zu mRNA / Protein  
(iii) HYPOTHETISCH: nein  
(iv) ANTISENSE: nein  
(v) ART DES FRAGMENTS: Gesamtsequenz  
5 (vi) URSPRÜNGLICHE HERKUNFT:  
(A) ORGANISMUS: *Cladosporium herbarum*  
(C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

DNA Sequenz 1323 b.p. ATGCCTATCTCC ... ATCAACTTGTA linear

```

10 1 / 1
ATG CCT ATC TCC AAG ATC CAC TCC CGC TAC GTC TAC GAC TCC CGT GGA AAC CCC ACC GTC
met pro ile ser lys ile his ser arg tyr val tyr asp ser arg gly asn pro thr val
61 / 21
GAG GTC GAC ATT GTC ACC GAG ACC GGT CTT CAC CGC GCC ATT GTC CCC TCT GTT GCC TCT
glu val asp ile val thr glu thr gly leu his arg ala ile val pro ser val ala ser
121 / 41
ACC GGC TCC CAC GAG GCT TGC GAG CTC CGT GAT GGT GAC AAG AGC AAG TGG GCT GGC AAG
thr gly ser his glu ala cys glu leu arg asp gly asp lys ser lys trp ala gly lys
181 / 61
GGT GTG ACC AAG GCT GTT GCC AAC GTC AAC GAG ATC ATT GCT CCC GCT CTC ATC AAG GAG
gly val thr lys ala val ala asn val asn glu ile ile ala pro ala leu ile lys glu
15 241 / 81
AAC CTC GAC GTC AAG GAC CAG GCC GCC GTT GAC GCT TTC CTC AAC AAG CTC GAT GGC ACC
asn leu asp val lys asp gln ala ala val asp ala phe leu asn lys leu asp gly thr

301 / 101
ACC AAC AAG ACC AAG ATT GGT GCC AAC GCC ATC CTC GGT GTC TCC ATG GCT GTG GCC AAG
thr asn lys thr lys ile gly ala asn ala ile leu gly val ser met ala val ala lys
361 / 121
GCT GCC GCC GCC GAG AAG CGT GTC CCT CTC TAC GCC CAC ATC AGC GAC CTT TCC GGC ACC
ala ala ala ala glu lys arg val pro leu tyr ala his ile ser asp leu ser gly thr
421 / 141
AAG AAG CCC TTC GTT CTC CCC GTT CCC TTC ATG AAC GTC GTC AAC GGT GGC TCC CAC GCg
lys lys pro phe val leu pro val pro phe met asn val val asn gly gly ser his ala
20 481 / 161
GGT GGC CGT CTT GCC TTC CAG GAG TTC ATG ATC GTT CCC AGC Ggg GCT CCC TCC TTC ACC
gly gly arg leu ala phe gln glu phe met ile val pro ser gly ala pro ser phe thr
541 / 181
GAG GCC ATG CGC CAG GGT GCT GAG GTC TAC CAG AAG CTG AAG TCC CTC ACC AAG AAG AGG
glu ala met arg gln gly ala glu val tyr gln lys leu lys ser leu thr lys lys arg
601 / 201
TAC GGC CAG TCT GCC GGC AAC GTC GGT GAC GAG GGT GGT GTC GCT CCC GAT ATT CAG ACC
tyr gly gln ser ala gly asn val gly asp glu gly gly val ala pro asp ile gln thr
661 / 221
GCT GAG GAG GCT CTT GAC CTC ATC ACC GAC GCC ATT GAG GAG GCC GGC TAC ACT GGC CAG
ala glu glu ala leu asp leu ile thr asp ala ile glu glu ala gly tyr thr gly gln
25 721 / 241
ATC AAG ATC GCC ATG GAC gtt GCC TCC TCC GAG TTC TAC AAG GCC GAC GAG AAG AAG TAC
ile lys ile ala met asp val ala ser ser glu phe tyr lys ala asp glu lys lys tyr
781 / 261
GAC CTT GAC TTC AAG AAC CCC GAC TCT GAC AAG AGC AAG TGG ATC ACC TAC GAG CAG CTT
asp leu asp phe lys asn pro asp ser asp lys ser lys trp ile thr tyr glu gln leu
841 / 281
GCC GAC CAG TAC AAC GAG CTC GCT GCC AAG TAC CCC ATT GTC TCC ATT GAG GAC CCC TTC
ala asp gln tyr asn glu leu ala ala lys tyr pro ile val ser ile glu asp pro phe
901 / 301
GCT GAG GAT GAC TGG GAG GCC TGG AGC TAC TTC TAC AAG ACC TCT GGC TCT GAC TTC CAG
ala glu asp asp trp glu ala trp ser tyr phe tyr lys thr ser gly ser asp phe gln
30 961 / 321
ATC GTCGGT GAT GAC CTT ACC GTC ACC AAC CCC GAG TTC ATC AAG AAG GCC ATC GAG ACC
ile val gly asp asp leu thr val thr asn pro glu phe ile lys lys ala ile glu thr

```

```

1021 / 341      1051 / 351
AAG GCC TGC AAC GCC CTC CTC CTC AAG GTC AAC CAG ATC GGT ACC ATC ACC GAG GCC ATC
lys ala cys asn ala leu leu leu lys val asn gln ile gly thr ile thr glu ala ile
1081 / 361      1111 / 371
AAC GCC GCC AAG GAC TCC TTC GCC GCT GGC TGG GGT GTC ATG GTC Tct cAc CGT TCC GGT
asn ala ala lys asp ser phe ala ala gly trp gly val met val ser his arg ser gly
1141 / 381      1171 / 391
GAG ACC GAG GAC GTC ACC atc gcc gac ATC GTT GTC GGT CTC CGT GCC GGC CAG ATc AAG
glu thr glu asp val thr ile ala asp ile val val gly leu arg ala gly gln ile lys
1201 / 401      1231 / 411
ACC GGT GCC CCT GCC CGT TCC GAG CGT CTC GCC AAG CTC AAC CAG ATC CTC CGT ATC GAG
thr gly ala pro ala arg ser glu arg leu ala lys leu asn gln ile leu arg ile glu
1261 / 421      1291 / 431
GAG GAG CTC GGT GAC AAG AGG CTC TAC GCC GGT GAC AAC TTC cgc ACT GCC ATC AAC TTG
glu glu leu gly asp lys arg leu tyr ala gly asp asn phe arg thr ala ile asn leu
1321 / 441
TAA
OCH

```

Nachfolgende Sequenzvergleiche (Seq.6) der Aminosäuresequenz mit der 10 SWISSPROT Proteinbank ergaben, daß Clah47 signifikante Homologie zu Enolasen zeigt. Im nachfolgenden "multiple sequence alignment" sieht man die hohen Homologien und Identitäten zwischen Clah47 und den übrigen Enolasen (Mensch, Ratte, Maus, Drosophila, Hefe).

#### Sequenz 6:

```

15  pileup.msf(enog_human)      1      50
    pileup.msf(enog_rat)      .SIqKIWARE IILSRGNPTV EVDLyTaKGL FRAAVPSGAS TGIYEALRLR
    pileup.msf(enog_mouse)    .SIqKIWARE IILSRGNPTV EVDLhTaKGL FRAAVPSGAS TGIYEALRLR
    pileup.msf(eno_drome)     .SIeKIWARE IILSRGNPTV EVDLyTaKGL FRAAVPSGAS TGIYEALRLR
    pileup.msf(eno1_yeast)    MtIkaIkARq IYDSRGNPTV EVDLTTElGL FRAAVPSGAS TGVHEALELR
    pileup.msf(eno2_yeast)    .aVSKVIARs VYDSRGNPTV EVELTTEKGV FRsIVPSGAS TGVHEALEMR
    pileup.msf(Eno_clado)     MpISKIhsRy VYDSRGNPTV EVDIVTetGL hRAIVPSVAS TGsHEACELR
    Consensus                -----R-- --DSRGNPTV EV---T--G- -R--VPS-AS TG--EA-E-R

20  pileup.msf(enog_human)      51      100
    pileup.msf(enog_rat)      DGDkQrYLgK GVLKAVDHIN stIAPALIss gLsVVEQEkI DNLMLELDGT
    pileup.msf(enog_mouse)    DGDkQrYLgK GVLKAVDHIN stIAPALIss gLsVVEQEkI DNLMLELDGT
    pileup.msf(eno_drome)     DnDKanYhGK sVLKAVgHVn srIAPALIss gIsVVEQEkI DNLMLELDGT
    pileup.msf(eno1_yeast)    DGDkSkWMGK GVLhAVKnVN dvIAPAFVka NLDVVDQkaV DdFLIsLDGT
    pileup.msf(eno2_yeast)    DeDKSkWMGK GVMhAVnnVN nvIAaAFVka NLDVVDQkaV DdFLIsLDGT
    pileup.msf(Eno_clado)     DGDkSkWagK GvTKAVanVN eIIAPALIKe NLDVVDQaaV DaFLnKLDGT
    Consensus                D-DK-----GK -V--AV---N ----- -V--Q--- D-----LDGT

25  pileup.msf(enog_human)      101      150
    pileup.msf(enog_rat)      ENKSKFGANA ILGVSLAVcK AGAAErElPL YRHIAqLAGN ..SDLiLPVP
    pileup.msf(enog_mouse)    ENKSKFGANA ILGVSLAVcK AGAAEKdIPL YRHIAqLAGN ..SDLiLPVP
    pileup.msf(eno_drome)     ENKSKFGANA ILGVSLAVaK AGAAKqVPL YkHIADLAGN ..kElILPVP
    pileup.msf(eno1_yeast)    aNKSILGANA ILGVSLAasr AaaaEKnVPL YkHIADLskS KTSpYVLPVP
    pileup.msf(eno2_yeast)    aNKSILGANA ILGVSHAAar AaaaEKnVPL YqHLADLskS KTSpYVLPVP
    pileup.msf(Eno_clado)     tNktKIGANA ILGVSHAVaK AaaaEKvPL YaHIsDLsGt K.kpFVLPVP
    Consensus                -NK-K-GANA ILGVs-A--- A-AA-----PL Y-H---L--- -----LPVP

30  pileup.msf(enog_human)      151      200
    pileup.msf(enog_rat)      AFNVINGGSH AGNKLAMQEF MiLPvGAESF RDAhRLGAeV YhTLKgVikd
    pileup.msf(enog_mouse)    AFNVINGGSH AGNKLAMQEF MiLPvGAESF RDAhRLGAeV YhTLKgVikd
    pileup.msf(eno_drome)     AFNVINGGSH AGNKLAMQEF MiLPtGATsF tEAMqMGsEV YhHLKnVikA
    pileup.msf(eno1_yeast)    fLNvINGGSH AGgaLALQEF MiAptGaktF aEALRIGsEV YhNLKsItKk
    pileup.msf(eno2_yeast)    fLNvINGGSH AGgaLALQEF MiAptGaktF aEAMRIGsEV YhNLKsItKk
    pileup.msf(Eno_clado)     fMNVVNGGSH AGgrLAFQEF MiVpSgAPsF tEAMRqGAeV YqKLKsItKk

```

```

Consensus  --NV-NGGSH AG--LA-QEF MI-P-GA--F --A---G-EV Y--LK---X-

201                                     250
pileup.msf(enog_human) KYGKDAtNVG DEGGFAPNII eNsEALeLVK EAIDKAGYTe KmvIGMDVAA
pileup.msf(enog_rat)   KYGKDAtNVG DEGGFAPNII eNsEALeLVK EAIDKAGYTe KmvIGMDVAA
pileup.msf(enog_mouse) KYGKDAtNVG DEGGFAPNII eNsEALeLVK EAIDKAGYTe KmvIGMDVAA
pileup.msf(eno_drome)  KFGlDATaVG DEGGFAPNIQ sNkEALnLIs DAIAKAGYTG KIEIGMDVAA
pileup.msf(eno1_yeast) rYGasAgNVG DEGGVAPNIQ taeEALDLIv DAIAaAGhdG KVKIGLDCAs
5 pileup.msf(eno2_yeast) rYGasAgNVG DEGGVAPNIQ taeEALDLIv DAIAaAGhdG KVKIGLDCAs
pileup.msf(Eno_clado)  rYGasAgNVG DEGGVAPNIQ taeEALDLIt DAIEeAGYTG qYkIaMDVAS
Consensus              --G--A--VG DEGG-AP-I- ---EAL-L-- -AI--AG--- ---I--D-A-

251                                     300
pileup.msf(enog_human) SEFYrDG..K YDLDFKs.Pt DpSrYITGDQ LgaLYQDFVr dYFVVSIEDP
pileup.msf(enog_rat)   SEFYrDG..K YDLDFKs.Pa DpSrCITGDQ LgaLYQDFVr nYFVVSIEDP
pileup.msf(enog_mouse) SEFYrDG..K YDLDFKs.Pa DpSrYITGDQ LgaLYQDFVr nYFVVSIEDP
pileup.msf(eno_drome)  SEFYKDG..q YDLDFKNeKs DKSqWlpaDk LAnLYkEFik dFPIVSIEDP
pileup.msf(eno1_yeast) SEFFKDG..K YDLDFKNPns DKSxWlTGpQ LAdLYhsLmk rYPIVSIEDP
pileup.msf(eno2_yeast) SEFFKDG..K YDLDFKNPes DKSxWlTGve LAdMYhsLmk rYPIVSIEDP
10 pileup.msf(Eno_clado) SEFYKadeK YDLDFKNPds DKSxWlTyEQ LAdqYnELaa kYPIVSIEDP
Consensus              SEF----- YDLDFK---- D-S----- L---Y----- --P-VSIEDP

301                                     350
pileup.msf(enog_human) FDQDDWaAWS KF..TanVGI QIVGDDLTvt NPKRIerAVE eKACNCLLLK
pileup.msf(enog_rat)   FDQDDWaAWS KF..TanVGI QIVGDDLTvt NPKRIerAVE eKACNCLLLK
pileup.msf(enog_mouse) FDQDDWaAWS KF..TanVGI QIVGDDLTvt NPKRIerAVE eKACNCLLLK
pileup.msf(eno_drome)  FDQDhWEAWS nL..TgctdI QIVGDDLTvt NPKRIatAVE kKACNCLLLK
pileup.msf(eno1_yeast) FaEDDWEAWS hFFKtag..I QIVaDDLTvt NPKRIataIE kKAadaLLLK
pileup.msf(eno2_yeast) FaEDDWEAWS hFFKtag..I QIVaDDLTvt NPaRIataIE kKAadaLLLK
pileup.msf(Eno_clado)  FaEDDWEAWS yFYKtsgsdf QIVGDDLTvt NPefIkkaIE tKACNaLLLK
Consensus              F--D-W-AWS ----T----- QIV-DDLTvt NP--I--A-E -KA---LLLK

351                                     400
15 pileup.msf(enog_human) VNQIGsVTEA IQackLAqen GWGVMVSHRS GETEDTFIAD LVVGLcTGQI
pileup.msf(enog_rat)   VNQIGsVTEA IQackLAqen GWGVMVSHRS GETEDTFIAD LVVGLcTGQI
pileup.msf(enog_mouse) VNQIGsVTEA IQackLAqen GWGVMVSHRS GETEDTFIAD LVVGLcTGQI
pileup.msf(eno_drome)  VNQIGTVTEs IaAhILakKn GWGtMVSHRS GETEDsFIgD LVVGLsTGQI
pileup.msf(eno1_yeast) VNQIGTIsEs IkaAqdsfaa GWGVMVSHRS GETEDTFIAD LVVGLrTGQI
pileup.msf(eno2_yeast) VNQIGTIsEs IkaAqdsfaa nWGVMVSHRS GETEDTFIAD LVVGLrTGQI
pileup.msf(Eno_clado)  VNQIGTITEA InAaKdsfaa GWGVMVSHRS GETEDvtIAD lVVGLraGQI
Consensus              VNQIG---E- I-A----- -WG-MVSHRS GETED--I-D -VVGL--GQI

401                                     442
pileup.msf(enog_human) KTGAPCRSER LAKYNQlMRI EEELGDEARF AGhNFRNPsv L.
pileup.msf(enog_rat)   KTGAPCRSER LAKYNQlMRI EEELGDEARF AGhNFRNPsv L.
20 pileup.msf(enog_mouse) KTGAPCRSER LAKYNQlMRI EEELGDEARF AGhNFRNPsv L.
pileup.msf(eno_drome)  KTGAPCRSER LAKYNQlMRI EEELGDEARF AGhNFRNPsv L.
pileup.msf(eno1_yeast) KTGAPaRSER LAKINQlLRI EEELGdNaVf AGeNFhhgdk L.
pileup.msf(eno2_yeast) KTGAPaRSER LAKINQlLRI EEELGdKAVY AGeNFhhgdk L.
pileup.msf(Eno_clado)  KTGAPaRSER LAKINQlLRI EEELGdkrly AGqNFRtain Lz
Consensus              KTGAP-RSER LAK-NQ--RI EEE-G----- AG--F----- --

```

Wie schon bei Clah53 erwähnt, kann auch in diesem Fall die Verbindung zwischen Proteinfunktion und Allergenität nicht hergestellt werden. Enolasen sind  
 25 aber aus einem anderen Punkt heraus interessant. Bei *Saccharomyces cerevisiae* konnte die Enolase (ENO1) als "heat-shock-Protein" nachgewiesen werden. Die Expression von ENO1 wird hier als Mehrgebrauch von Energie in dieser schwierigen Stressbedingung gedeutet (Iida und Yahara 1985). In Hefe wird Enolase auch in der stationären Phase sowie unter Schwefelhungern vermehrt exprimiert (Cohen 1987).

Die nachfolgende Sequenz 7 zeigt mit Computerunterstützung gefundene B-Zellepitope. Hoher Antigenindex, unter Berücksichtigung von Sekundärstruktur, Oberflächenlage, Hydrophilität, Flexibilität etc.

- 5 Sequenz 7: Clah47=Enolase\_clado: B-Zellepitope
- (1) ANGABEN ZU SEQ ID NO:7
  - (i) SEQUENZKENNZEICHEN:
    - (A) LÄNGE: einzeln angeführt
    - (B) ART: Protein
    - (ii) ART DES MOLEKÜLS: Peptide
  - 10 (iii) HYPOTHETISCH: nein
  - (v) ART DES FRAGMENTS: N-Terminus bis C-Terminus
  - (vi) URSPRÜNGLICHE HERKUNFT:
    - (A) ORGANISMUS: Cladosporium herbarum
    - (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen
- 15 Tyr Val Tyr Asp Ser Leu Gly Asn Pro Thr Val Glu Val (10-22)
- Pro Ser Val Ala Ser Thr Gly Ser His Glu Ala Cys Glu Leu Arg Asp Gly Asp Lys Ser Lys Trp
- Ala Gly Lys Gly Val Thr Lys (36-64)
- Asn Lys Leu Asp Gly Thr Thr Asn Lys Thr Lys Ile Gly Ala (95-108)
- Ala His Ile Ser Asp Leu Ser Gly Thr Lys Lys Pro Phe Val (132-145)
- Tyr Gln Lys Leu Lys Ser Leu Thr Lys Lys Arg Tyr Gly Gln Ser Ala Gly Asn Val Gly Asp Glu
- 20 (190-211)
- Phe Lys Asn Pro Asp Ser Asp Lys Ser Lys Trp Ile Thr Tyr Glu (264-278)
- Ser Tyr Phe Tyr Lys Thr Ser Gly Ser Asp Phe (309-319)
- Ala Ile Asn Ala Ala Lys Asp Ser Phe Ser Ala Gly Trp Gly (359-372)
- Met Val Ser His Arg Ser Gly Glu Thr Glu Asp Val Thr Met (374-387)
- His Ile Val Arg Arg Ser Arg Ala Gly Gln Ile Lys Thr Gly (388-401)
- 25 Ala Pro Ala Arg Ser Asp Gly Leu Ala Lys Leu Asn (402-413)
- Ile Leu Arg Ile Glu Glu Glu Leu Gly Asp Lys Arg Leu Tyr Ala Gly Asp Asn Phe Arg Thr Ala
- (415-436)

Die nachfolgende Sequenz 8 zeigt die berechneten T-Zellepitope im 30 1-Lettercode. Amphipathische Bereiche mit einem Score geringer als 10 wurden für nicht relevant angenommen.

# 5 Sequenz 8: Vorausgesagte amphipathatische Segment

## T-Zellepitope

IHSRYVYDSLGN  
 KGVTKAVANVNEIAP  
 DAFLNKLDGT  
 10 AHISDLSG  
 PSFTEAMRQGAEVYQKLKSLTK  
 GQSAGNV  
 EALDLITDAIEE  
 LADQYNEL  
 EAWSYFY  
 VNQIGTITEAINAAK  
 DGLAKLNQILRIE

15

(1) ANGABEN ZU SEQ ID NO:8

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: einzeln angeführt

20 (B) ART: Protein

(ii) ART DES MOLEKÜLS: Peptide

(iii) HYPOTHETISCH: nein

(v) ART DES FRAGMENTS: N-Terminus bis C-Terminus

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Cladosporium herbarum

25 (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

Ile His Ser Arg Tyr Val Tyr Asp Ser Leu Gly Asn (6-17)

Lys Gly Val Thr Lys Ala Val Ala Asn Val Asn Glu Ile Ile Ala Pro (60-75)

Asp Ala Phe Leu Asn Lys Leu Asp Gly Thr (91-100)

Ala His Ile Ser Asp Leu Ser Gly (132-139)

30 Pro Ser Phe Thr Glu Ala Met Arg Gln Gly Ala Glu Val Tyr Gln Lys Leu Lys Ser Leu Thr Lys  
(177-198)

Gly Gln Ser Ala Gly Asn Val (202-208)  
 Glu Ala Leu Asp Leu Ile Thr Asp Ala Ile Glu Glu (223-234)  
 Leu Ala Asp Gln Tyr Asn Glu Leu (280-287)  
 Glu Ala Trp Ser Tyr Phe Tyr (306-312)  
 Val Asn Gln Ile Gly Thr Ile Thr Glu Ala Ile Asn Ala Ala Lys (350-364)  
 5 Asp Gly Leu Ala Lys Leu Asn Gln Ile Leu Arg Ile Glu Glu (407-420)

Die T-Zellepitope errechnen sich aus den Aminosäurepositionen der Midpoints, die N-terminal von einem Lysin (K), C-terminal von einem Prolin (P) flankiert werden (=Flags). Es sind nur dann potentielle T-Zellepitope vorhanden, wenn der "Score-Index" größer als 10 ist.

10

### C. Clah22

Die nachfolgende Sequenz 9 zeigt die vollständige cDNA Sequenz von Clah22. Die daraus abgeleitete Aminosäuresequenz ist ebenfalls ersichtlich. Der offene Leserahmen von Clah22 beträgt 615bp, was einer Aminosäurelänge von 205  
 15 Aminosäuren entspricht. Das berechnete Molekulargewicht des rekombinanten Proteins beträgt 22341 Dalton.

Sequenz 9: YCP4\_clado -> 1-phasen Translation 22341 Dalton

(1) ANGABEN ZU SEQ ID NO:9

20

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 615 Basenpaare / 205 Aminosäurereste

(B) ART: Nukleinsäure / Protein

(C) STRANGFORM: ds

(D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: cDNA zu mRNA / Protein

(iii) HYPOTHETISCH: nein

(iv) ANTISENSE: nein

(v) ART DES FRAGMENTS: Gesamtsequenz

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Cladosporium herbarum

30 (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

DNA sequence 615 b.p. ATGGCTCCCAAG ... AACTTCCAGTGA linear

```

1   /   1
ATG GCT CCC AAG ATC GCC ATC ATC TTC TAC TCC ACC TGG GGA CAC GTC CAG ACC CTC GCC
met ala pro lys ile ala ile ile phe tyr ser thr trp gly his val gln thr leu ala
61 / 22
5 GAG GCC GAG GGC AAG GGC ATC CGC GAG GCC GGC GGT TCC GTC GAC CTC TAC CGC GTC CCC
glu ala glu ala lys gly ile arg glu ala gly gly ser val asp leu tyr arg val pro
122 / 41
GAG ACC CTC ACC CAG GAG GTT CTG ACC AAG ATG CAC GCT CCT CCT AAG GAC GAC TCC ATC
glu thr leu thr gln glu val leu thr lys met his ala pro pro lys asp asp ser ile
181 / 61
CCT GAG ATC ACC GAC CCT TTC ATT CTC GAG CAG TAC GAT CGC TTT CCT CAT GGG CAT CCC
pro glu ile thr asp pro phe ile leu glu gln tyr asp arg phe pro his gly his pro
241 / 81
ACC CGC TAC GGC AAC TTC CCC GCG CAG TGG AGG ACC TTC TGG GAC CGC ACC GGC GGC CAA
thr arg tyr gly asn phe pro ala gln trp arg thr phe trp asp arg thr gly gly gln
301 / 101
10 TGG CAG ACC GGT GCC TTC TGG GGC AAG TAC GCC GGC CTG TTC ATC AGT ACT GGT ACC CAG
trp gln thr gly ala phe trp gly lys tyr ala gly leu phe ile ser thr gly thr gln
361 / 122
GGT GGT GGA CAG GAG AGC ACT GCT CTC GCT GCG ATG AGC ACT CTC TCC CAC CAC GGC ATT
gly gly gly gln glu ser thr ala leu ala ala met ser thr leu ser his his gly ile
422 / 141
ATC TAC GTT CCT CTT GGC TAC AAG ACC ACA TTC CAC CTC CTC GGC GAC AAC AGC GAG GTC
ile tyr val pro leu gly tyr lys thr thr phe his leu leu gly asp asn ser glu val
481 / 161
CGC GGT GCC GCA GTG TGG GGC GCT GGT ACC TTC TCC GGT GGT GAC GGT TCC CGC CAG CCC
arg gly ala ala val trp gly ala gly thr phe ser gly gly asp gly ser arg gln pro
541 / 181
15 TCC CAG AAG GAA TTG GAG CTC ACC GCC CAG GGT AAG GCC TTC TAT GAG GCT GTC GCC AAG
ser gln lys glu leu glu leu thr ala gln gly lys ala phe tyr glu ala val ala lys
601 / 201
GTC AAC TTC CAG TGA
val asn phe gln opa

```

Homologiesuchen mit dem sequenzierten Protein in der SWISSPROT-Proteindatenbank zeigten, daß das Allergen Clah22 signifikante 20 Homologie zu dem Hefeprotein YCP4 aufweist. Die Identität der beiden Proteine beträgt 56%, die Homologie sogar 70%. Eine so hohe Ähnlichkeit läßt wohl eine gemeinsame Funktion dieser beiden Proteine vermuten. Die nachfolgende Sequenz 10 spiegelt die hohe Homologie von Clah22 und YCP4 wieder.

#### Sequenz 10:

25 ycp4\_yeast x ycp4\_clado

```

1 MV.KIAIITYSTYGHIDVLAQAVIKGVZAAGGKADIYRVEETLPDEVLT 49
1. ||||| |||:|:|:|:|:| ||: .|||..|:|:|:|:|:|
1 MAPKIAIIFYSTWGHVQTLAEAEAKGIREAGGSVDLYRVPETLTQEVLT 50
30 50 MNAPOKPEDIPVATEKLTLE.YDAFLFGVPTFRFNGNLPQWSAFWDKTGGL 98
1:|:|:|:|:|:| 1: .:|:| || | | ||:|:|:|:|:|:|
51 MHAPPKDDSIPEITDFFILEQYDRFPHGHPTRYGNFPAQWRFTWDRGTGGQ 100
99 WAKGSLNGKAAGIFVSTSSYGGGQESTVKACLSYLAHHGIIFLPLGYKNS 148

```

```

101 WQTGAFWGYAGLFISTGTQGGGQESTALAAMSTLSHHGIIYVPLGYKTT 150
149 FAELASIEZEVHGGSPWGAGTLAGPDGSRRTASPLELRIAEIQGKTFYETAK 198
151 FHLLGDNSEVRGAAVWGAGTFSGGDGSRQPSQKELETA.QGKAFYEAVA 199
199 KL.FPA 203
200 KVNFGZ 205

```

5

Die Sequenz, bzw. der offene Leserahmen von YCP4, wurde im Rahmen des Hefegenomprojektes am Chromosom 3 von *Saccharomyces cerevisiae* lokalisiert und publiziert (Biteau et al. 1992). Eine durchgeführte Disruption von YCP4 zeigte nach (Biteau et al. 1992) keinen Phänotyp.

Es hat sich auch gezeigt, daß auch Clah22 einen homologen Partner in *Alternaria alternata* besitzt. Die folgende Sequenz 11 zeigt ein "sequence alignment" zwischen den Allergenen Alta22 und Clah22.

Sequenz 11:

15 ycp4\_alt x ycp4\_clado

```

1 MAPKIAIVYYSMYGHIMADAEKGIQEAGGDAKLFQVAETLPQEVLDK 50
101 WQTGAFWGYAGLFISTGTQGGGQESTALAAMSTLSHHGIIYVPLGYKTT 150
1 MAPKIAIIFYSTWGHVQTLAEAEAKGIREAGGSVDLYRVPETLTQEVLT 50
101 WQTGAFWGYAGLFISTGTQGGGQESTALAAMSTLSHHGIIYVPLGYKTT 150
51 MYAPPKDSSVPVLEDPVLEEFDGILFGIPTRYGNFPAQFKTFWDKTGKQ 100
101 WQTGAFWGYAGLFISTGTQGGGQESTALAAMSTLSHHGIIYVPLGYKTT 150
51 MHAPPKDDSIPEITDPPFILEQYDRFPHGHPTRYGNFPAQWRTFWDRGTGGQ 100
101 WQTGAFWGYAGLFISTGTQGGGQESTALAAMSTLSHHGIIYVPLGYKTT 150
101 WQTGAFWGYAGLFISTGTQGGGQESTALAAMSTLSHHGIIYVPLGYKTT 150
151 FSLANLDEVHGGSPWGAGTFSAGDGSRQPS..ELELNIAQAQKAFYE 198
151 FHLLGDNSEVRGAAVWGAGTFSGGDGSRQPSQKELETA..AQKAFYE 197
199 VAKAH.QZ 205
198 VAKVNFQZ 205

```

25 Die mit Computerunterstützung gefundenen B-Zellepitope sind in der nächsten Sequenz 12 zu sehen.

Sequenz 12: Clah22=YCP4\_clado: B-Zellepitope

(1) ANGABEN ZU SEQ ID NO:12

30

(i) SEQUENZKENNZEICHEN:



- (A) LÄNGE: einzeln angeführt
- (B) ART: Protein
- (ii) ART DES MOLEKÜLS: Peptide
- (iii) HYPOTHETISCH: nein
- (v) ART DES FRAGMENTS: N-Terminus bis C-Terminus
- 5 (vi) URSPRÜNGLICHE HERKUNFT:
  - (A) ORGANISMUS: Cladosporium herbarum
  - (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

Tyr Ser Thr Trp Gly His Val Gln (10-17)  
 Glu Ala Glu Ala Lys Gly Ile Arg Glu Ala Gly Gly Ser (21-33)  
 10 Lys Met His Ala Pro Pro Lys Asp Asp Ser Ile Pro Glu Ile Thr Asp Pro (50-66)  
 Leu Glu Gln Tyr Asp Arg Phe Pro His Gly His Pro Thr Arg Tyr Gly Asn Phe (69-86)  
 Pro Ala Gln Trp Arg Thr Phe Trp Asp Arg Thr Gly Gly Gln Trp Gln Thr Gly (87-104)  
 Ile Ser Thr Gly Thr Gln Gly Gly Gly Gln Glu Ser Thr Ala Leu Ala (115-130)  
 Ile Tyr Val Pro Leu Gly Tyr Lys Thr Thr Phe (141-151)  
 Leu Leu Gly Asp Asn Ser Glu Val Arg Gly Ala (153-163)  
 15 Gly Ala Gly Thr Phe Ser Gly Gly Asp Gly Ser Arg Gln Pro Ser Gln Lys Glu Leu Glu Leu Thr  
 (167-188)

Die folgende Sequenz 13 zeigt die berechneten T-Zellepitope. Amphipathische Helices, flankiert von hydrophilen Bereichen stellen das Grundmuster der Berechnung für MHC II assoziierte Peptide dar.

20

Sequenz 13: Vorausgesagte amphipathatische Segmente

T-Zellepitope

YSTWGHVQTLAEA

IREA

25 GSVDLYRVPETLTQEVLTSMH

DSIPEITD

YDRFPHGHPTRYGNFPAQWRTFWDRTGGQ

AAMSTLS

AFYEA

(1) ANGABEN ZU SEQ ID NO: 13

30

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: einzeln angeführt
- (B) ART: Protein
- (ii) ART DES MOLEKÜLS: Peptide
- (iii) HYPOTHETISCH: nein
- (v) ART DES FRAGMENTS: N-Terminus bis C-Terminus
- 5 (vi) URSPRÜNGLICHE HERKUNFT:
  - (A) ORGANISMUS: Cladosporium herbarum
  - (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

Tyr Ser Thr Trp Gly His Val Gln Thr Leu Ala Glu Ala (10-22)

Ile Arg Glu Ala (27-30)

10 Gly Ser Val Asp Leu Tyr Arg Val Pro Glu Thr Leu Thr Gln Glu Val Leu Thr Lys Met His  
(32-52)

Asp Ser Ile Pro Glu Ile Thr Asp (58-65)

Tyr Asp Arg Phe Pro His Gly His Pro Thr Arg Tyr Gly Asn Phe Pro Ala Gln Trp Arg Thr Phe  
Trp Asp Arg Thr Gly Gly Gln (72-100)

Ala Ala Met Ser Thr Leu Ser (130-136)

15 Ala Phe Tyr Glu Ala (193-197)

Die T-Zellepitope errechnen sich aus den Aminosäurepositionen der Midpoints, die N-terminal von einem Lysin (K), C-terminal von einem Prolin (P) flankiert werden (=Flags). Es sind nur dann potentielle T-Zellepitope vorhanden, wenn der "Score-Index" größer als 10 ist.

20

#### D. Clah11

Die folgende Sequenz 14 zeigt die vollständige cDNA-Sequenz von Clah11 und die von ihr abgeleiteten Aminosäuresequenz. Der offene Leserahmen umfaßt 336bp bzw. 112 Aminosäuren. Das berechnete Molekulargewicht beträgt 11078  
25 Dalton und entspricht somit dem 11kD großen antigenen Protein, das im Westernblot von 4% der Patienten erkannt wird.

Sequenz 14: Clah11=rla2\_clado -> 1-phasen Translation 11078 Dalton

30

(1) ANGABEN ZU SEQ ID NO:14

- (i) SEQUENZKENNZEICHEN:  
 (A) LÄNGE: 336 Basenpaare / 112 Aminosäurereste  
 (B) ART: Nukleinsäure / Protein  
 (C) STRANGFORM: ds  
 5 (D) TOPOLOGIE: linear  
 (ii) ART DES MOLEKÜLS: cDNA zu mRNA / Protein  
 (iii) HYPOTHETISCH: nein  
 (iv) ANTISENSE: nein  
 (v) ART DES FRAGMENTS: Gesamtsequenz  
 (vi) URSPRÜNGLICHE HERKUNFT:  
 10 (A) ORGANISMUS: Cladosporium herbarum  
 (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

DNA Sequenz 336 b.p. ATGAAGTACCTC ... CTCTTCGACTAA linear

```

1   /   1   .....   31   /   11
ATG AAG TAC CTC GCA GCT TTC CTC CTC CTC GGC TTG CCG GGC AAC TCC TCC CCC TCT GCT
15 met lys tyr leu ala ala phe leu leu leu gly leu pro gly asn ser ser pro ser ala
61 / 22 91 / 31
GAG GAC ATC AAG ACC GTC CTC AGC TCC GTT GGC ATC GAC GCC GAC GAG GAG CCG TCT CAG
glu asp ile lys thr val leu ser ser val gly ile asp ala asp glu glu pro ser gln
122 / 41 151 / 51
CTC CTC CTT AAG GAG CTC GAG GGC AAG GAC ATC AAC GAG CTG ATC TCT TCC GGC TCT GAG
leu leu leu lys glu leu glu gly lys asp ile asn glu leu ile ser ser gly ser glu
181 / 61 221 / 71
AAG CTC GCC TCCGGTT CCC TCT GGT GGT GCT GGT GCC GCC TCT GCC GGT GGT GCT GCC GCC
lys leu ala ser val pro ser gly gly ala gly ala ala ser ala gly gly ala ala ala
241 / 81 271 / 91
GCT GGT GGT GCC GCT GAG GCC GCC CCC GAG GCC GAG AAG GCT GAG GAG GAG AAG GAG GAG
20 ala gly gly ala ala glu ala ala pro glu ala glu lys ala glu glu glu lys glu glu
301 / 101 331 / 111
TCT GAC GAC GAC ATG GGC TTC GGT CTC TTC GAC TAA
ser asp asp asp met gly phe gly leu phe asp OCH

```

- Sequenzvergleiche mittels FASTA in der SWISSPROT-Proteindatenbank ergaben, daß das vorliegende 11kD große allergen wirkende Protein signifikante Homologien bzw. Identitäten zu RLA2, einem hoch konservierten ribosomalen Protein, besitzt. Die nachfolgende Sequenz 15 spiegelt die hohe Homologie wieder. Gezeigt sind "multiple alignments" von Mensch, Ratte, Drosophila, Schizosaccharomyces pombe, Saccharomyces cerevisiae, Dictyostelium, Trypanosoma und den beiden Pilzen Cladosporium herbarum und Alternaria alternans.

### Sequenz 15:

```

30
pileup.msf(rla2_human) 1 50
pileup.msf(rla2_rat) MrYVAsYLLa aLGGNsSPSA kDIKKILDSV GIEADDDRln KVISELnGKn

```

```

pileup.msf(rla2_drome) MrYvAAYLLa VLGGkdSPan SDleKILsSV GVEvDaERlt KVlkeLaGks
pileup.msf(rla2_alt)  HQhLAAYLLl GLGGntSPSA aDVKaVLESV GIZADsDRlD KLISeLEGKO
pileup.msf(rla2_clado) MKYLAAYLLl GL.GNsSPSA eDIKtVLSV GIDADEEpsq 111kELEGKO
pileup.msf(rla2_schpo) MKYLAAYLLl tvGGkdSPSA SDiesVLstV GIEAEsERIE tlinELnGKO
pileup.msf(rla2_yeast) MKYLAAYLLl naaGnt.PdA tkIKaILESV GIEIEDEkvs sVLSaLEGKS
pileup.msf(rla2_dicdi) .KYLAAYLLa sLsGN..ana asVtKILqSV GVEvDaaRvE sVckELDgKO
pileup.msf(rla2_trycr) MKYLAAYaLv GLsGgt.PSk SaVeaVLkaa GvpvDpsRvD alfaEFaGKO
Consensus      ----A---L- ---G-----L---G-----GK-

51
5 pileup.msf(rla2_human) IEDvIAqGig KL.ASVPAGG AvAvSAapgs AApAAGsApa AAEEKKDEKK 100
pileup.msf(rla2_rat)    IEDvIAqGvg KL.ASVPAGG AvAvSAapgs AApAAGsApa AAEEKKDEKK
pileup.msf(rla2_drome) IDDLIkEGrE KL.sSmPVGG ggAvAA..d AApAAAaggD kxEaKKEEK
pileup.msf(rla2_alt)    InELIASGsE KL.ASVPsGG AggAAAsggA AAaggsAqaE AApEaake..
pileup.msf(rla2_clado) InELIssGsE KL.ASVPsGG AgAASAggAA AAgg.....AEKaeE..
pileup.msf(rla2_schpo) IDELIAAGnE KL.AtVptGG AaaA.ApAAA AggAApAAeE AAKeakeE..
pileup.msf(rla2_yeast) VDELItEGnE KL.AaVPAag pasAggAAAA sgdA.....AAEEeKEE..
pileup.msf(rla2_dicdi) VqalIAaGks Kv.gsVaAaa ApAAatsAap AAaAApAkk vvEEKK....
pileup.msf(rla2_trycr) fDtvctEGks KL.vggVtrpn AatASaptAA AAassGAAap AAaae.....
Consensus      -----G--K-----

101      117
10 pileup.msf(rla2_human) EESEESDDDM GFGLFD.
pileup.msf(rla2_rat)    EESEESDDDM GFGLFD.
pileup.msf(rla2_drome) EESEsEDDDM GFaLFZ.
pileup.msf(rla2_alt)    EEKESEDDDM GFGLFDZ
pileup.msf(rla2_clado) EkKEESDDDM GFGLFDZ
pileup.msf(rla2_schpo) E..EESDDDM GFGLFD.
pileup.msf(rla2_yeast) EaaEESDDDM GFGLFD.
pileup.msf(rla2_dicdi) ...EESDDDM GmGLFD.
pileup.msf(rla2_trycr) ...EEeDDDM GFGLFD.
Consensus      ---E--D-DM G--LF--

```

- 15 Das allergene Protein Clah11 ist nicht nur wegen seiner Eigenschaft als Allergen von *Cladosporium herbarum* interessant. Ribosomale Proteine, hier im speziellen die humanen ribosomalen Proteine P1 und P2, sind in der Literatur als Autoantigene beschrieben worden (Francoeur et al. 1985, Rich et al. 1987, Hines et al. 1991). 20% der Patienten mit Lupus erythematosus besitzen Autoantikörper (anti-rRNP) gegen Komponenten der Ribosomen, im speziellen Autoantikörper
- 20 gegen die ribosomalen Proteine P0 (38kD), P1 (16kD) und P2 (15kD). Das P2 Protein entspricht in seiner Homologie dem allergenen Protein Clah11. Die humanen Autoantikörper kreuzreagieren mit ähnlichen Proteinen, was heißt, daß Epitope erkannt werden die in der Evolution stark konserviert wurden. Die Basis der immunologischen Kreuzreaktivität bildet die 17 Aminosäurereste lange carboxyterminale Region KEESEESD(D/E)DMGFGLFD. Ob eine in der Kindheit
- 25 und Jugend erfolgte Sensibilisierung durch Clah11 mit einem im Erwachsenenalter auftretenden Autoimmunkrankheit korreliert, bedarf einer genauen Prüfung. Applikationen von ribosomalen Proteinen konnten allerdings in Mäusen keine Autoimmunkrankheit erzeugen (Hines et al. 1991).

Die gezeigten B-Zellepitope in der nächsten Sequenz 16 sind unter Berücksichtigung von Sekundärstruktur, Oberflächenlage, Hydrophilität, Flexibilität 30 etc. berechnet worden.

## Sequenz 16: Clah11=rla2\_clado: B-Zellepitope

## (1) ANGABEN ZU SEQ ID NO:16

## (i) SEQUENZKENNZEICHEN:

5 (A) LÄNGE: einzeln angeführt

(B) ART: Protein

(ii) ART DES MOLEKÜLS: Peptide

(iii) HYPOTHETISCH: nein

(v) ART DES FRAGMENTS: N-Terminus bis C-Terminus

(vi) URSPRÜNGLICHE HERKUNFT:

10 (A) ORGANISMUS: Cladosporium herbarum

(C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

Gly Leu Pro Gly Asn Ser Ser Pro Ser Ala Glu Asp Ile Lys Thr Val Leu (11-27)

Gly Ile Asp Ala Asp Glu Glu Pro Ser Gln Leu Leu Leu (31-43)

Lys Glu Leu Glu Gly Lys Asp Ile Asn Glu Leu (44-54)

15 Ser Ser Gly Ser Glu Lys Leu Ala Ser Val Pro Ser Gly Gly Ala Gly (56-71)

Ala Glu Ala Ala Pro Glu Ala Glu Lys Ala Glu Glu Glu Lys Glu Glu Ser Asp Asp Asp Met Gly  
Phe (85-107)

Die nachfolgende Sequenz 17 zeigt die berechneten T-Zellepitope. Bereiche mit einem Score geringer als 10 werden als nicht relevant angenommen.

20

## Sequenz 17: Vorausgesagte amphipathatische Segmente

T-Zellepitope

EDIKTV

EGKDINELISSGSEKLASVPSG

25 ASAGG

## (1) ANGABEN ZU SEQ ID NO:17

## (i) SEQUENZKENNZEICHEN:

30 (A) LÄNGE: einzeln angeführt

(B) ART: Protein

- (ii) ART DES MOLEKÜLS: Peptide
- (iii) HYPOTHETISCH: nein
- (v) ART DES FRAGMENTS: N-Terminus bis C-Terminus
- (vi) URSPRÜNGLICHE HERKUNFT:
  - (A) ORGANISMUS: Cladosporium herbarum
- 5 (C) ENTWICKLUNGSSTADIUM: Sporen und vegetative Hyphen

Glu Asp Ile Lys Thr Val (21-26)

Glu Gly Lys Asp Ile Asn Glu Leu Ile Ser Ser Gly Ser Glu Lys Leu Ala Ser Val Pro Ser Gly  
(47-68)

Ala Ser Ala Gly Gly (82-86)

10

Die T-Zellepitope errechnen sich aus den Aminosäurepositionen der Midpoints, die N-terminal von einem Lysin (K), C-terminal von einem Prolin (P) flankiert werden (=Flags). Es sind nur dann potentielle T-Zellepitope vorhanden, wenn der "Score-Index" größer als 10 ist.

15

20

25

30

## Literatur

- Aukrust, L. (1979).  
5 Cross radioimmuno-electrophoretic studies of distinct allergens in two extracts of *Cladosporium herbarum*.  
Int. Arch. Allergy Appl. Immunol. 58, 371.
- Aukrust, L. (1980).  
Allergens in *Cladosporium herbarum*.  
10 In *Advances in Allergology and Immunology*. Edited by A. Oehling. Oxford, Pergamon Press.
- Birkner, T., Rumpold, H., Jarolim, E., Ebner, H., Breitenbach, M., Skarvil, F., Scheiner, O., Kraft, D. (1990).  
Evaluation of immunotherapy-induced changes in specific IgE, IgG and IgG  
15 subclasses in birch pollen allergic patients by means of immunoblotting. Correlation with clinical response.  
Allergy 45, 418.
- Bold, H.C., Alexopoulos, C.J., Delevoryas, T. (1973).  
Morphology of plants and fungi.  
20 New York, Harper and Row.
- Cohen, R.T., Yokoi, J.P., Holland, A.E., Pepper, A.E., Holland, M.J. (1987).  
Transcription of the constitutively expressed yeast enolase gene ENO1 is mediated by positive and negative cis-acting regulatory sequences.  
Mol. Cell. Biol. 7, 2753.  
25
- Ferreira, F.D., Hoffmann-Sommergruber, K., Breiteneder, H., Pettenburger, K., Ebner, C., Sommergruber, W., Steiner, R., Bohle, B., Sperr, W.R., Valent, P., Kungl, A.J., Breitenbach, M., Kraft, D., Scheiner, O. (1993).  
Purification and characterization of recombinant BetvI, the major birch pollen allergen. Immunological equivalence to natural BetvI.  
30 J. Biol. Chem. in press.

Foucard, T. Dreborg, S., Sten, E. (1984).  
Mould Allergy Workshop. Uppsala, Sweden:  
Ord & Form; Pharmacia Diagnostics AB 1984.

Francoeur, A.M., Peebles, C.L., Heckman, K.J., Lee, J.C., Tan, E.M. (1985).  
5 Identification of ribosomal protein autoantigens.  
J. Immunol. 135, 1767.

Gell, P.G.H., Coombs, R.R.A., Lachmann, P.J. (1975).  
Clinical Aspects of Immunology.  
Blackwell, Oxford.

10

Gravesen, S. (1979).  
Fungi as a cause of allergic disease.  
Allergy 34, 135.

Harada, S., Agarwal, D.P., Goedde, H.W. (1982).  
15 Mechanism of alcohol sensitivity and disulfiram-ethanol reaction.  
Subst. Alco. Act. Misuse. 3, 107.

Hines, J.J., Weissbach, H., Brot, N., Elkon, K. (1991).  
Anti-P autoantibody production requires P1/P2 as immunogens but is not driven by  
exogenous self-antigen in mrl mice.  
20 J. Immunol.. 146, 3386.

Hsu, L.C., Bendel, R.E., Yoshida, A. (1987).  
Direct detection of usual and atypical alleles on the human aldehyde dehydrogenase-2  
(ALDH2) locus.  
Am. J. Hum. Genet. 41, 996.

25

Iida, H., Yahara, i. (1985).  
Yeast heat shock protein of MW 48000 is an isoprotein of enolase.  
Nature 315, 688.

Lacey, J. (1981).  
30 The aerobiology of conidial fungi.  
In Biology of conidial fungi. Vol 1.



New York, Academic Press.

Margalit, H., Spogue, J.L., Cornette, J.L., Cease, K.B., Delisi, C., Berzofsky, J.A. (1987).

Prediction of immunodominant Helper T cell antigenic sites from the primary  
5 sequence.

J. Immunol. 138, 2213.

Rammensee, H.G., Falk, K., Rötzschke, O. (1993).

MHC molecules as peptide receptors.

Current Opinion in Immunol. 5, 35.

10

Rich, B.E., Steitz, J.A. (1987).

Human acidic ribosomal phosphoproteins P0, P1 and P2: analysis of cDNA clones,  
in vitro synthesis and assembly.

15 Mol. Cell. Biol. 7, 4065.

Rothbard, J.B., Gefter, M.L. (1991).

Interactions between immunogenic peptides and MHC proteins.

Ann. Rev. Immunol. 9, 527.

20 Sanger, F., Nicklen, S., Coulson, A.R. (1977).

DNA sequencing with chain-terminating inhibitors.

Proc. Natl. Acad. Sci. USA 74, 5463-5468

25

30

ERSATZBLATT (REGEL 26)

## Patentansprüche

1. Rekombinante DNA Moleküle, die für Polypeptide kodieren, die die  
5 Antigenität der Allergene Clah53, Clah47, Clah22 und Clah11 besitzen oder für  
Peptide, die mindestens ein Epitop dieser Allergene aufweisen, dadurch  
gekennzeichnet, daß sie Nukleinsäuresequenzen aufweisen, die mit den Sequenzen  
1, 3-5, 7-9, 12-14 sowie 16 und 17, oder mit Teilbereichen dieser Sequenzen in  
homologer Weise übereinstimmen, bzw. Nucleinsäuresequenzen, die mit den  
genannten Nucleinsäuresequenzen unter stringenten Bedingungen hybridisieren.
- 10 2. Rekombinante DNA-Moleküle nach Anspruch 1, dadurch gekennzeichnet,  
daß sie Nukleinsäuresequenzen aufweisen, die durch Degeneration aus den  
dargestellten Sequenzen 1, 3-5, 7-9, 12-14 sowie 16 und 17 ableitbar sind.
3. Rekombinante DNA Moleküle nach Anspruch 1 oder 2, dadurch  
gekennzeichnet, daß sie Nukleinsäuresequenzen aufweisen, die für Polypeptide  
kodieren, die als Antigene kreuzreaktiv mit den Allergenen Clah53, Clah47, Clah22  
15 und Clah11 sind und zu diesen eine hohe Homologie aufweisen.
4. Rekombinante DNA-Moleküle nach Ansprüchen 1 bis 3 dadurch  
gekennzeichnet, daß sie funktionell mit einer Expressionskontrollsequenz zu einem  
Expressionskonstrukt verbunden sind.
5. Wirtssystem zur Expression von Polypeptiden, dadurch gekennzeichnet,  
daß es mit einem rekombinanten Expressionskonstrukt nach Anspruch 4 transformiert  
20 ist.
6. Aus einem DNA-Molekül nach einem der Ansprüche 1 bis 3 abgeleitetes  
rekombinantes oder synthetisches Protein oder Polypeptid, dadurch gekennzeichnet,  
daß es die Antigenität von Clah53, Clah47, Clah22 oder Clah11, oder zumindest von  
einem Epitop dieser Proteine, aufweist.
7. Rekombinantes oder synthetisches Protein oder ein Polypeptid nach  
25 Anspruch 6, dadurch gekennzeichnet, daß es eine Aminosäuresequenz aufweist, die  
den gezeigten Sequenzen 1, 3-5, 7-9, 12-14 sowie 16 und 17 zur Gänze oder  
teilweise entspricht.
8. Rekombinantes oder synthetisches Protein oder Polypeptid nach Anspruch  
6 oder 7, dadurch gekennzeichnet, daß es ein Fusionsprodukt darstellt, das die  
Antigenität der Allergene Clah53, Clah47, Clah22 oder Clah11, oder zumindest  
30 eines Epitops davon aufweisen und einen zusätzlichen Polypeptidanteil besitzt, wobei

das gesamte Fusionsprodukt von der DNA eines Expressionskonstrukts gemäß Anspruch 4 kodiert wird.

9. Rekombinantes oder synthetisches Protein oder Polypeptid nach Anspruch 5 8, dadurch gekennzeichnet, daß der besagte zusätzliche Polypeptidanteil  $\beta$ -Galaktosidase oder ein anderes zur Fusion geeignetes Polypeptid ist.

10. Diagnostisches oder therapeutisches Reagens, dadurch gekennzeichnet, daß es ein synthetisches Protein oder Polypeptid gemäß einem der Ansprüche 6 bis 9 enthält.

11. Verfahren zum in vitro -Nachweis der Allergie eines Patienten gegen die 10 Allergene Clah53, Clah47, Clah22 oder Clah11, dadurch gekennzeichnet, daß die Reaktion der IgE Antikörper im Serum des Patienten mit einem rekombinanten oder synthetischen Protein oder Polypeptid nach einem der Ansprüche 6 bis 9 gemessen wird.

12. Verfahren zum in vitro - Nachweis der zellulären Reaktion auf die Allergene Clah53, Clah47, Clah22 oder Clah11, dadurch gekennzeichnet, daß ein 15 rekombinantes oder synthetisches Protein oder Polypeptid nach einem der Ansprüche 6 bis 9 zur Stimulierung oder Hemmung der zellulären Reaktion eingesetzt wird.

20

25

30

Fig.1: Westernblotting eines 13,5%iges Polyacrylamidgels, nach Auftrennung von *Cladosporium herbarum* Proteinextrakt und Inkubation mit Sera verschiedener Patienten.

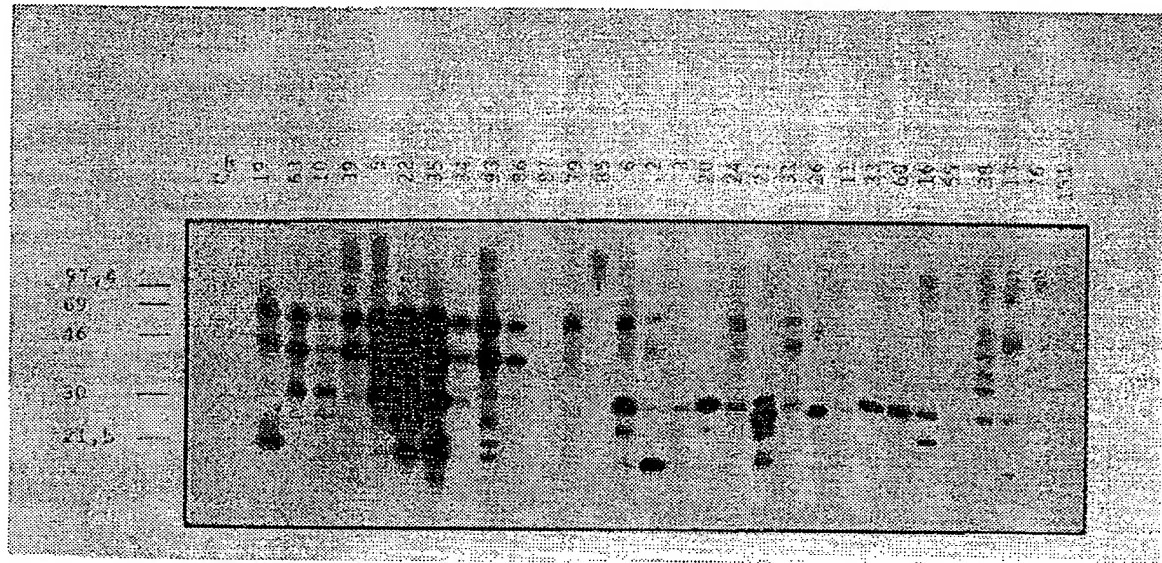
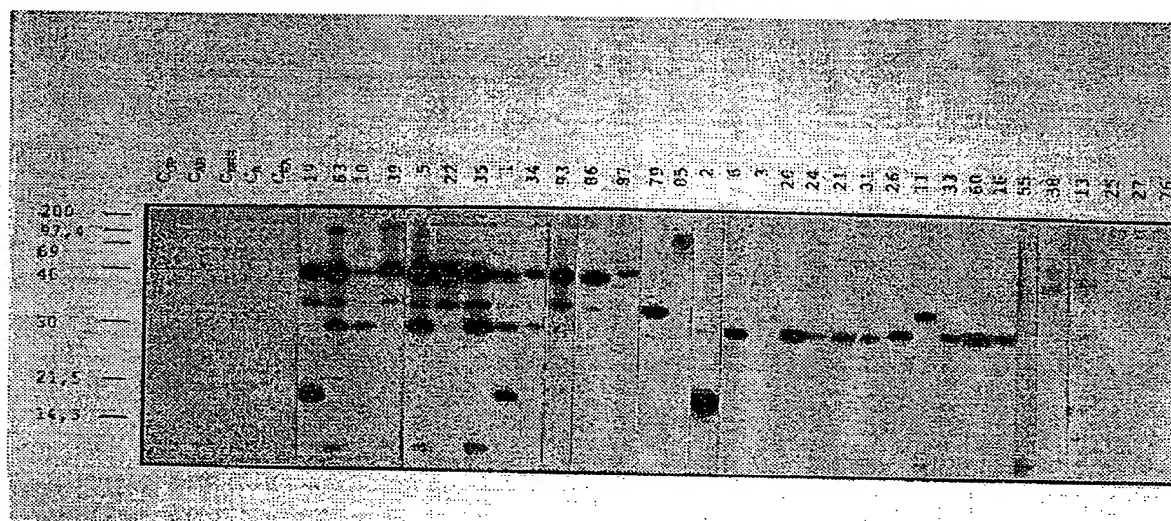
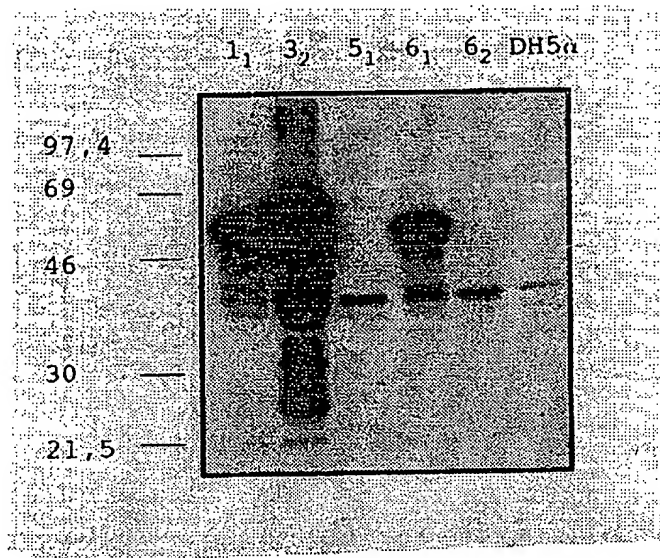


Fig.2: Auftrennung von *Cladosporium herbarum* Proteinextrakt auf einem 17,5%igen Polyacrylamidgel; Inkubation mit Patientensera; Detektion mit Jod-markiertem anti-human IgE



2/2

Fig.3: Expression der rekombinanten Proteine Clah47 und Clah53 in Bluescript nach IPTG Induktion.



**THIS PAGE BLANK (USPTO)**